A discusión

SELECTION AND EFFICIENCY IN HIERARCHICAL SOCIAL SYSTEMS*

José A. García Martínez

WP-AD 2005-35

Correspondence: Departamento de Fundamentos del Análisis Económico, Universidad de Alicante, Campus de Sant Vicent del Raspeig, Apartado de Correos 99, 03080 Alicante, Spain.

Editor: Instituto Valenciano de Investigaciones Económicas, S.A.
Primera Edición Diciembre 2005
Depósito Legal: V-98-2006

* I wish to express my gratitude to Fernando Vega Redondo, my thesis advisor, for his encouragement and invaluable comments. I also thank Pablo Beker, Giovanni Ponti, Joseph E. Harrington, Francisco Marhuenda, Frédéric Palomino and an anonymous referee for their comments and Jose Bordallo for helping with graphics. Financial support from the Instituto Valenciano de Investigaciones Económicas (IVIE) is gratefully acknowledged.

SELECTION AND EFFICIENCY IN HIERARCHICAL SOCIAL SYSTEMS

José A. García Martínez

ABSTRACT

We examine the influence that the degree of stringency in the promotion processes of hierarchical systems has on the outcome of such selections at both the local and global level. We show that any change in the degree of stringency, whether an increase or a decrease, could cause counterintuitive effects. In our analysis, we consider a hierarchical system in which there is a large population of agents at each level. Specifically, this level is a continuum of two different kinds of agents, one with a greater expected performance (i.e., a higher expected success rate) than the other. The agents interact among themselves in groups and in an environment that is stochastic and idiosyncratic (for the group). The social institution promotes agents a posteriori, on the basis of their performance. We consider diverse systems for such promotion processes, with varying degrees of stringency, and study their long-run behaviour patterns.

We find that the degree of stringency can be counter-productive for the homogeneity of a population in hierarchical systems. In the process for the selection of agents to be promoted, an increase in stringency beyond a certain point favours, surprisingly enough, agents with lower probabilities of success. Thus, the more stringent the system is, the more heterogeneous the population becomes in the long run, both kinds of agents survives. On the other hand, when the stringency of the system is maintained between two fixed thresholds, the more-successful agents are the only ones who survive. Finally, when the stringency of the system is too low, both homogeneous equilibria (we only consider two kinds of agents) are stable and the dynamics depend on the initial conditions, two basin of attraction appear.

Keywords: Hierarchies, Stringency-Toughness Selection Degree, Evolution.

1 Introduction.

Selection processes exist in all societies. In general terms, agents interact among themselves, within organizations and institutions, and as a result of such interaction, certain individuals are promoted over others, achieving a higher status in the form of greater political or economic power, increased prestige, more responsibility and a wider intellectual influence. In such selection processes, the characteristics of the individual agents obviously play an important role, but institutional factors like the degree stringency or a hierarchical structure, also exert a great influence on the selection process.

The aim of our study is to understand the role that certain institutional factors play in deciding the outcome of social selection processes in the long-run.

To clarify the concept of the degree of stringency, we consider the following scenario: There is a set of agents who may, but might not, interact among themselves. As such, we may rank them by outcomes or by any other "agent's characteristic' we consider to be of interest. We then select a fraction $\alpha \in (0, 1)$ of the highest ranks. The closer α is to zero, the greater the stringency in the selection process, and the closer it is to one, the less stringent the selection is. In effect, in the latter case, there is practically no "selection' process. We could define the degree of stringency in the promotion system as $1 - \alpha$, i.e., one minus the fraction of agents promoted.

We examine the influence that the degree of stringency in the promotion processes of hierarchical systems has on the outcome of such selections at both the local and global level. We show that any change in the degree of stringency, whether an increase or a decrease, could cause counterintuitive effects.

We shall now present a brief review of the most pertinent literature on the topic and the relationship with our work.

Harrington[e.g., 1998a,1998b] focuses on the flexibility of the agents' behavior, (rigidity vs flexibility). Agents are organized and interact within an evolving social hierarchy with different levels, like a pyramid. The agents are all randomly picked and paired off with other agents of their own levels, against whom they must compete. The successful agent is then promoted to the next level. In such a bilateral competition, the agents must respond correctly to an exogenous and stochastic environment, which is idiosyncratic to any pair of agents. There are, in fact, two right responses, $\{A, B\}$. Harrington considers three kinds of rules on behavior, two of which are rigid and a third which is flexible. A rigid rule always triggers Action A (B), irrespective of the environment. The flexible rule, on the other hand, always allows the right response to the current environment, and agents can modify their actions as need be. Furthermore, as the agents all respond to the same rule while they remain in the system, we can then talk about rigid and flexible agents. Harrington finds that rigid agents do better in the long run than flexible agents do, eventually being the only survivors in his social system. A crucial assumption in Harrington[eg., 1998a,1998b] to arrive at this outcome is that a change in a flexible agent's action is accompanied by a reduction of his capacity to implement the action.

In Harrington[1999a,1999b], he follows his line of study but now introduces the concept of "social learning", i.e., young agents, who observe the elder ones at the top of the hierarchy and imitate them in a bid to become as successful as they are. Imitation, however, is not perfection. His results in this case are quite similar to those of his previous analysis: i.e., in the long run, the followers of the rigid rule are the most successful. In Harrington's model there seems to be only local selection. One agent competes with another of his own level and one of them is always subsequently promoted, regardless of what might be happening within the rest of their level. As such, half of the population of each level is always promoted in each period.

In our model, we link the degree of stringency in the promotion system to the number or fraction of agents it promotes. The higher the number of people promoted, the less stringent the system is, and, inversely, the fewer promoted, the more stringent it is. As the system automatically promotes first the agents who have the best outcomes (in their pairs), the more stringent the system is, the more difficult promotion will be for the agents with poorer outcomes.

In Harrington's model, we could consider the degree of stringency to be always equal to $\frac{1}{2}$. Half of the population that competes is promoted. In other words, one agent out of every two, in each period. However, there is no reason why a greater number of agents could not compete in groups. Let us say, for instance, that instead of pairs, a group of n agents compete together.

If just one agent out of every five is promoted, the system is obviously more stringent than it would be if five out of six were. The former is a more stringent model than Harrington's and the latter is much less stringent than his is.

One of the main objectives of this study is to discover the role that the degree of local stringency plays in the evolution of such hierarchies.

In many social systems, however, there is not only a local competition for promotion, but also a broader, more global competition throughout each level. In the academic world, for example, students first compete at their own universities (i.e., the local competition), and once they have finished their courses, the most successful ones from all of the different universities, (i.e., the best outcomes of the local competitions), then compete for a place in a university department, (i.e., the global competition for that level). If they are successful enough to get a place, they then compete locally again, within their new departments, in an effort to obtain a Ph.D. degree. After achieving the Ph.D., however, they must then compete globally once more, with other Ph.D. graduates from all of the other universities, for a limited number of jobs in a given number of organizations. Similar processes exist among people who work in a firm. They first compete within their own groups or departments (locally) and the best ones are candidates to be promoted The successful candidates, however, must then compete again at a more global level, with other successful people from other departments, for promotion to a higher status.

As such, we can consider the selection process as a two-step development. First, there is a local competition, which produces a more or less intense local selection. The intensity of this first selection depends on the peculiar institutional parameters that govern the selection, like the degree of local stringency for example. The second stage is a broader or more global competition, among the agents who were selected in the first round The final selection, therefore, depends (among others things) on both the local and the global degrees of stringency in the selection processes.

Vega-Redondo[2000] employs this kind of global selection, using a hierarchical system that is quite similar to Harrington's, although with a different approach and purpose. In each period, two agents of the same level are randomly matched in a bilateral coordination game, as in Harrington's model. As they must follow the same strategy while they remain in the system, we can consider two kinds of agents, those who follow the efficient (equilibrium) strategy and those who choose the other one. There are four possible outcomes, as expressed below. A proportion ρ of them is promoted to the next level. The agents with the best payoffs are promoted first, but the system continues selecting agents until all of the vacancies (determined by the parameter ρ) have been filled. After the local competition, agents are ordered in four classes, according to their payoffs. The agents with the best payoffs are those who have followed the efficient strategy and were matched against other agents who also followed the efficient strategy. The two next best payoffs are for the agents who chose the inefficient strategy, and the worst payoffs are for those who followed the efficient strategy but were matched against others who chose the inefficient strategy. As such, the agents who followed the efficient strategy are either at the top of at the bottom of the outcome ranking.

Vega-Redondo shows that if the parameter ρ is lower than a certain threshold, the unique long-run state induces all of the candidates to play the efficient strategy. On the other hand, if the promotion is rather lax and ρ is greater than the fixed threshold, the unique long-run state induces all of the candidates to play the inefficient strategy. An increase in the stringency degree, therefore, favours long-run efficiency, since it increases the fragility of the inefficient configuration.

We should therefore emphasize the fact that while the payoffs in both Harrington's and Vega-Redondo's models are obtained in local competitions, there is no local selection in Vega-Redondo's model and the local selection in Harrington's model is fixed and always equal to $\frac{1}{2}$.

Our principal objective, as we have already explained, is to understand the influence that the degree of stringency (local and global) in the promotion processes of hierarchical systems have on the outcome of such selections. More specifically, we wish to find out what happens when stringency is increased. Does the behavior of the general population become more homogeneous or more heterogeneous as a result of an increase? Is an increase in stringency favorable to agents with greater expected performances? And finally, is an increase in local stringency equivalent or comparable to one in global stringency?

In our search for answers, we consider a hierarchy in a setting which, for the sake of simplicity, is quite similar to Harrington's model. In other words, there is no interaction of strategies and the agents' payoffs are independent of the population's profile. We then parameterize the degrees of both the local and the global stringencies separately and quite easily. Only two kinds of agents are considered, one type being more successful in expected terms than the other and, therefore, having a relative advantage. The agents interact randomly in groups of n individuals, with each agent of a given level being matched with another $n-1$ agents of his own level, after which their outcomes, good or bad, are evaluated.

After obtaining the outcomes for all of the agents, the selection system works in two separate stages: It first selects the k (with $k \leq n$) agents with the best outcomes from each group of n agents, (i.e., the local selection[1]) . The second step is the global selection, when the agents who were chosen in the local selection are now compared according to their previous outcomes and a fraction θ of them with the best outcomes are chosen. As such, only the agents who have been selected in both the local and the global selections are promoted to the next level.

In summary then, we have a family of promotion systems with different degrees of stringency. $S_{[k,n,\theta]}$ is a selection system that works in a hierarchy in which the agents are randomly matched in groups of n agents. In its first step, the system selects k individuals from each group of n agents, and in the second step, a proportion θ of the agents thus selected are finally promoted. Therefore, $\frac{k}{n}$ will be the local promotion degree and determines the degree of local stringency, and θ will be the global promotion degree and determines the degree of global stringency. In the promotion system of Harrington's model, for instance, $k=1$, $n=2$, $\theta=1$.

The present paper is also related with the literature of tournaments developed since the seminal work of Lazear and Rosen (1981). In this literature we find a similar

[1] The selected agents are a known proportion $\frac{k}{n}$ of all of the agents who have been matched.

organization structure as our hierarchical model, for example in Rosen (1986). The promotion of agents is based on the relative performance of the contestant, as in our model. However, this literature is mainly focused on incentive problems. The tournament is considered as an incentive device, where the strategic variables of the agents are effort, and in some papers this variable is risk. (And this literature tries to answer question like what must be the optimal prize structure to achieve the optimal level of effort). The tournament literature has mostly focused on the case with homogeneous agents, where selection problems in the sense discussed here do not arise. Nevertheless, there are some works which have focused on selection role of contests, e.g., Rosen (1986, Section V), which consider both the motivation and the selection function of a contest and Hvide and Kristiansen(2002).

In Rosen (1981) the agents can influence in his probability of promotion by choosing a level of effort. There are several types of agents with different level of ability. For the same level of effort, the more ability an agent has, the greater his promotion probability is. Rosen consider only a degree of stringency in his hierarchical system. Thus, if we fix the effort level (so eliminate the strategic interaction) we obtain one of the different possible specification of our model. Unlike Rosen, we are interested in study the role of the stringency of promotion in the evolution of the hierarchy.

In Hvide and Kristiansen (2002). The agents can influence in his probability of promotion by choosing a level of risk. There are two types of agents, the high and low type. These authors firstly consider that the agents compete in pairs, however, he also consider the effect of an increase in the number of contestants. Thus, they also consider different degrees of stringency in the promotion. They obtain the result that an increase in the degree of stringency could have inefficient outcome, for them inefficiency is a decrease in the promotion probability of the agents of type high. The reason for this result is the strategic interaction. An increase in the number of contestants increases the probability that some contestants can be type high, in that case, all the agents have incentive to assume more risk, and this fact could decrease the probability of promotion for type high agents. As Hvide and Kristiansen, we also find that an increase in the degree of stringency could have an inefficient outcome. However, as we do not consider strategic interaction, there must be other reason for this inefficiency, as we will see in the present paper. Thus, we show that even in the case without strategic interaction the increase of the stringency could have inefficient outcomes.

Our model also present other difference with these papers of tournaments, unlike them we use a dynamic approach.

To conclude this introduction, we will say that the main result of our work is that the degree of stringency can be counter-productive for the homogeneity of a population in hierarchical systems. An increase in stringency beyond a certain point causes the population of agents with the highest expected success rate to decrease in the long run.

The reason for this is that an increase in the subpopulation of agents who follow a certain rule could have either a positive or a negative effects on the dynamics of this subpopulation. The final sign of this effect depends on the values of the institutional factors, like the degree of stringency in the selection process, and on the expected success rates assigned to the agents. In other words, the larger the population of a given type of agent, the greater their probability of achieving promotion will depend on institutional parameters, and the less it will depend on their success rates. This is because the larger the population of a given kind of agent is, (near homogeneity, we suppose), the greater the probability that his opponents will be of his own type, (a probability close to one), and will therefore be following the same behavior rule as him. As such, they will be either all successful or all unsuccessful. In such a case, the probability of promotion depends mainly on the degree of stringency in the selection process, (i.e. how many people are promoted). In contrast, the smaller the population of a given kind of agent is, the more their probability of promotion will depend on their success rates, and the less it will depend on institutional parameters. In such a case, if the population is sparse and close to extinction, an individual from that population will almost certainly be matched (with a probability close to one) against other agents with a different behavior rule. Thus, if he is successful, he will be the only successful agent in his group and, as such, his probability of promotion will depend mainly on his success rate. In the conclusions of this study, we analyze the crucial assumption to be able to arrive at this result.

The remainder of this article is presented as follows: Section 2 describes the model, and Sections 3, 4 and 5 analyze the problem. In Section 3 , we concentrate on the case of $n = 2$ and $k = 1$, in Section 4, for any n and $k = 1$, and in Section 5, the general case. In Section 6 we study the optimal population profile in accordance with the different institutional payoffs that we consider.

2 The Model.

Consider a hierarchical system with several different levels. At each level there is a large population of agents, specifically, a continuum of agents. We normalize the size of the level population to one. For the agents to carry out activities in the system, they follow one of two strategies or behavioral rules, either A or B. An agent follows the same strategy as long as he remains in the system. In each period, agents (from the same level) are randomly matched into groups to do a certain activity. The numbers of agents in each group is n. The activity is carried out in a stochastic environment (which is idiosyncratic for the groups), and the agents, therefore, face an environment of uncertain future.

We consider just two possible outcomes from the agents' activities, either right or wrong [2].

We can categorize all of the different possible environments into four types:

- Type 1: Strategy A is the right response to the environment and Strategy B is wrong.
- Type 2: Strategy B is the right response to the environment and the strategy A is wrong.
- Type 3: Strategies A and B both are wrong responses to the environment.
- Type 4: Strategies A and B both are right responses to the environment.

The probability of an environment of type i is P_i and $\sum_{i=1}^{4} P_i = 1$ with $i = 1, 2, 3, 4$. As we consider that $P_4 = 0$, we suppose that both strategies cannot be right in any given environment [3]. We label the environment of type i as e_i, with $i = 1, 2, 3$.

At each level, therefore, a proportion P_1 (P_2, P_3) of the set of groups has a type-1 $(2, 3)$ environment. This is assumed to be i.i.d. across levels, so that the probability that an agent faces a given environment is independent of the environment he has faced in the past. Although there is individual uncertainty as to the environment, the absence of aggregate uncertainty simplifies matters.

[2] We suppose that the system cannot evaluate how good the right activity is or how bad the activity wrong is.

[3] This assumption does not affect the results qualitatively.

An agent who follows Strategy A (B) has a probability of doing the right action equal to P_1 (P_2); we could consider this probability as a measure of the "expected" performance of this agent. If $P_1 > P_2$, then the agents who follow Rule A have the highest expected success rate. When we will say "success rate" of a rule, we will refer to the "expected success rate" of following this rule. It holds in all sections of the paper.

Once the agents have carried out their activities in a specific environment, the selection system begins and some of the agents are promoted to the next level. This selection is based on the "a posteriori" performance of the agents, i.e., the outcomes of their local interactions in their own groups.

The concept is quite similar to the way sports leagues are organized. There are different levels and, at each level, players compete in separate groups and the best player in each group is promoted to the next level. The final goal of each player is to get the top of the pyramid.

2.1 Promotion Mechanism

In each period, the population of a given level[4] is randomly matched in groups of n agents, who then implement their behavior rules in an idiosyncratic environment. The outcome is either success or failure in the activity.

The promotion mechanism then begins. It has two distinct steps:

First step, *Local Selection*: The k agents ($k \geq 1$) with the best outcomes are chosen from each group of n individuals as candidates for promotion, they are the "Eligible Agents". If there are more agents with the same outcome than available vacancies, they are chosen randomly to fill the available places. The Eligible Agents represent $\frac{k}{n}$ of the population of that particular level.

Second step, *Global Selection*: a proportion θ of the Eligible Agents is selected. The system first selects the Eligible Agents who were the successful competitors in local competition, i.e., they gave the right response to the their environments, and if all of the vacancies are not filled, it then chooses randomly from the Eligible Agents who failed in their local competitions, i.e., they did not give the right response to the their environments. The agents eventually promoted represent $\theta\frac{k}{n}$ of the population of that particular level, (which has been normalized to one), see Figure 1

Both the Local and the Global selections, therefore, are based on the outcomes of the competing agents. The agents with the best outcomes are always selected first.

We call the parameter θ the global degree of promotion and it measures the **stringency degree**[5] in the global selection. If $\theta = 1$, there is not global selection, as in the Harrington's model, in which case, all the Eligible Agents are promoted to the next level.

We call the ratio $\frac{k}{n}$ the local degree of promotion and it measures the degree of stringency in the local selection. If $\frac{k}{n} = 1$, then there is not local selection, as in Vega-Redondo's model, in which case the entire population of agents in the level concerned become Eligible Agents.

[4] Which has been normalized to one.

[5] We parametrize the **stringency degree** with the proportion of the population that was promoted. In fact, it is one minus the proportion of promoted agents. As the system first promotes the agents with the best outcomes, the more stringent it is, the more difficult it becomes for the agents with worst outcomes to be promoted. In other words, if there are 100 vacancies for promotion, for example, and only the 5 best candidates are chosen, the selection process is obviously much more stringent than it would be if it promoted 50 candidates. Since the proportion of promoted agents in the local selection is $\frac{k}{n}$ and θ in the global selection, $\frac{k}{n}\theta$ will be the absolute degree of promotion and $1 - \frac{k}{n}\theta$ the absolute degree of stringency.

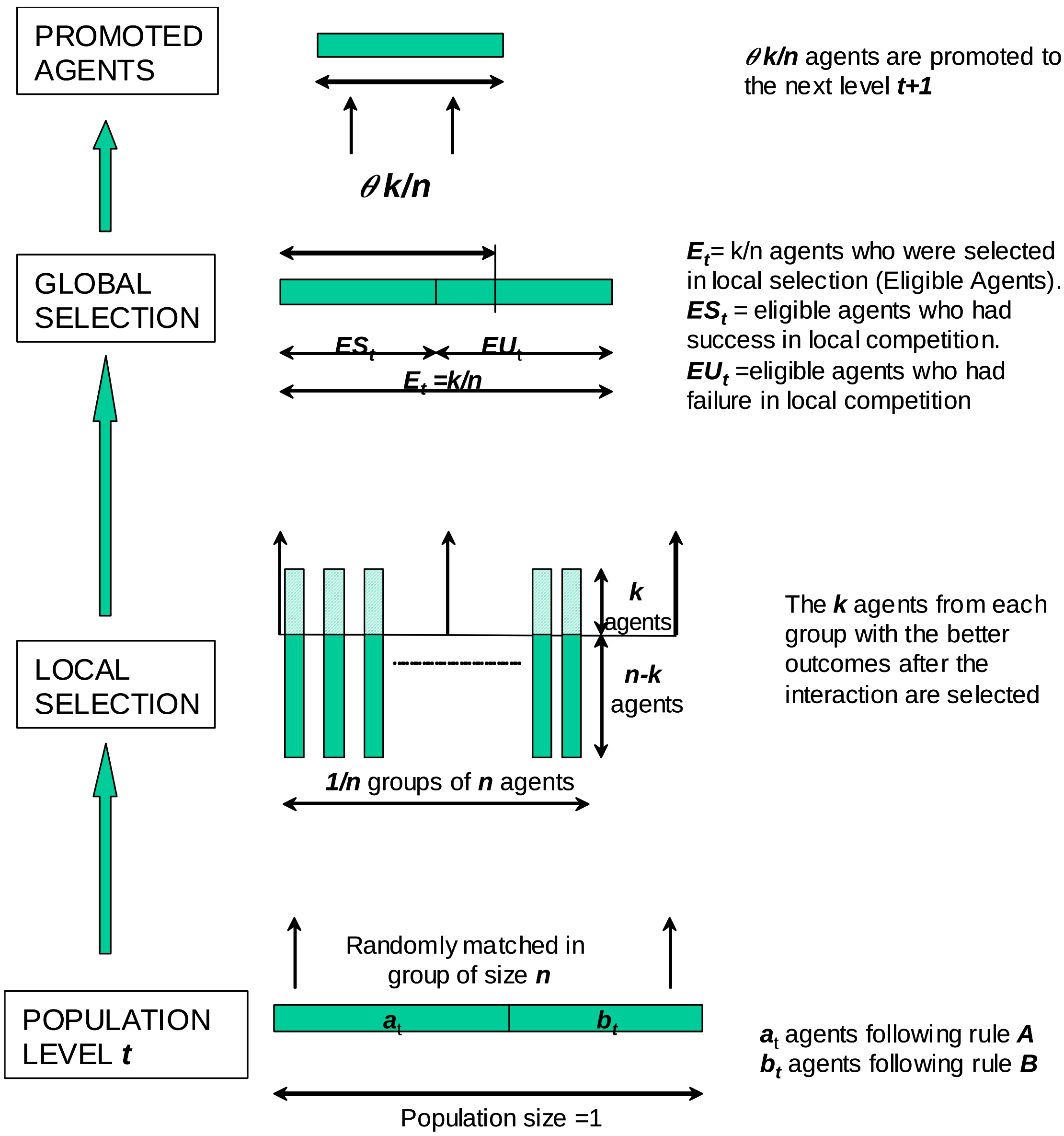


Figure 1: The promotion mechanism.

By $S_{[n,k,\theta]}$ we denote the promotion system that selects k agents from groups of n agents, in the local selection, and a proportion θ of Eligible Agents in the global selection. $S_{[2,1,1]}$ would therefore be the promotion system used in Harrington's model, and $S_{[2,2,\theta]}$ the one employed in the Vega-Redondo model. He, however, models the interaction among agents quite differently from either Harrington or us.

We follow both Harrington[1998a,1999a] and Vega-Redondo[2000] in making the simplifying assumption that, at each t, only agents who have not yet suffered a set-back in their promotion are considered to be "still in the race", and the rest are no longer to be considered for promotion, i.e. it is an up or out system.

2.2 The dynamics

The dynamic system functions as follows: The entire initial population lives at the lowest social level and is comprised of agents who are endowed with either rule A or B. They are match up in groups, in an idiosyncratic environment, after which, the selection system begins and some of them are selected for promotion to the next level. In the following period, the process is repeated, but only among the agents who were selected in the first round, up or out..

If the hierarchy is to be kept going "full" a fresh cohort of agents must enter the lowest level at the end of each period to replace the agents who have moved up a rung It is assumed that the younger agents, who come into the system, imitate the older successful agents at the top of the ladder. Note that to consider such hierarchical system is the same as considering a hierarchical system with a lower limit but without any upper bound, and we study what happens to a single cohort which evolves up through the system. In the long run, the two systems behave alike and provide the same selection. We decided, therefore, to study what happens to a single cohort in a hierarchy with no upper bound. The specific objective is to observe the evolution of the characteristics of a cohort of agents as it migrates upward through the hierarchy.

Let a_t denote the proportion of the population of level t who follow Strategy A, and b_t denotes Strategy B (Note that $b_t = 1 - a_t$). An agent who follows Strategy A (B) is referred to as an A-Type (B -Type) agent. As such, we need only study the dynamics of a_t, since the population that follows Strategy B will be $b_t = 1 - a_t$. Finally, let n be the number of agents in a group.

In short, the state of the system is given by $a_t \in [0, 1]$. The probabilities of exogenous environments is given by $P_1, P_2 \in (0, 1); P_1 + P_2 \leq 1$, and the institutional parameters are $\theta \in (0, 1]$ and $k > 1$.

Definition 1 $a^* \in [0, 1]$ *is a globally stable equilibrium of the dynamic system given by* $a_t = f(a_{t-1})$ *if for all* $a_0 \in (0, 1)$ $\lim_{t \to \infty} a_t = a^*$

$a^*[X]$ denotes the globally stable equilibrium under the selection process X.

3 Selection in Pairwise Contest.

In this section, we study the case in which $n = 2$, i.e., one agent is matched with another agent. Thus, the local competition is in groups of two agents. This is enough to understand the role of the parameter θ in the long-run selection. Obviously, in such a case the parameter k can only have two values; either $k = 1$ or $k = 2$. In the latter there is no local selection. Therefore, if $n = 2$, there are only two degrees of stringency in the local selection.

In the first place, we study the case in which $k = 1$. In other words, we study the family of the promotion system $S_{[n=2,k=1,\theta]}$. The second part of this section addresses the case $S_{[n=2,k=2,\theta]}$

3.1 The promotion system $S[n = 2, k = 1, \theta]$.

We first show how we obtain the dynamic equation.

In the first period, the agents are matched in pairs, after which, we have three possible kinds of groups; (AA) both agents following A, (BB) both agents following B and (AB) each agent following a different rule. Since we normalize the population of a given level to one, the number of groups[6] is $\frac{1}{2}$, since the size of each group is two $(n = 2)$. At level t, and after the random matching, there are $\frac{1}{2}a_t^2$ groups[7] AA, $a_t\, b_t$ groups AB, and $\frac{1}{2}b_t^2$ groups BB:

kind of group	AA	AB	BB
number of groups	$\frac{1}{2}a_t^2$	$a_t b_t$	$\frac{1}{2}b_t^2$

The eligible[8] agents (E_t) are one from each group. The number of eligible agents who were successful in the activity is [9]:

$$ES_t = \frac{1}{2}a_t^2\, P_1 + a_t\, b_t(\, P_1 + P_2) + \frac{1}{2}b_t^2\, P_2$$

The successful eligible agents who follow Strategy A are therefore:

$$ES_t^a = \frac{1}{2}a_t^2\, P_1 + a_t\, b_t\, P_1$$

The eligible agents who were unsuccessful are[10]:

$$EU_t = \frac{1}{2}a_t^2(P_2 + P_3) + \frac{1}{2}b_t^2(P_1 + P_3) + a_t\, b_t\, P_3$$

The unsuccessful eligible agents who follow Strategy A are:

$$EU_t^a = \frac{1}{2}a_t^2\,(P_2 + P_3) + \frac{1}{2}\,a_t\, b_t\, P_3$$

It must be remembered that the system selects a proportion θ of the eligible agents (E_t). It first selects the successful eligible agents (ES_t) and if there are not enough of

[6] The number of groups is equal to the number of individual in the population divided by two

[7] As the population is infinite, we can identify probabilities with frequencies or proportions.

[8] Note that the eligible (E_t) agents are $\frac{1}{2}$ of the population, $\frac{1}{2} = E_t = ES_t + EU_t$

[9] We have $\frac{1}{2}a_t^2$ eligible agents who were in AA groups, (one agent from each group). Moreover, a proportion P_1 of them were the environment e_1 and, therefore, were successful. There are also $a_t b_t$ agents from the AB groups, and a proportion $P_1 + P_2$ were in e_1 or e_2 and were successful as well. And, a proportion P_2 of the agents from the BB groups were successful as well.

[10] Note that $EU_t = \dfrac{1}{2} - ES_t$.

them, it then selects from the unsuccessful eligible agents (EU_t). Since $E_t = \frac{1}{2}$, the system selects $\frac{1}{2}\theta$ eligible agents.

Before obtaining the dynamic equation, we must consider two different possible cases:

First case. If $\theta\frac{1}{2} \leq ES_t$, i.e., the number of agents ($\theta\frac{1}{2}$) selected from the eligible ones is smaller than or equal to the eligible agents that are successful (ES), then, the system only selects agents who were successful. In Figure 2 we can see all the eligible agents (1/2 of the total population).

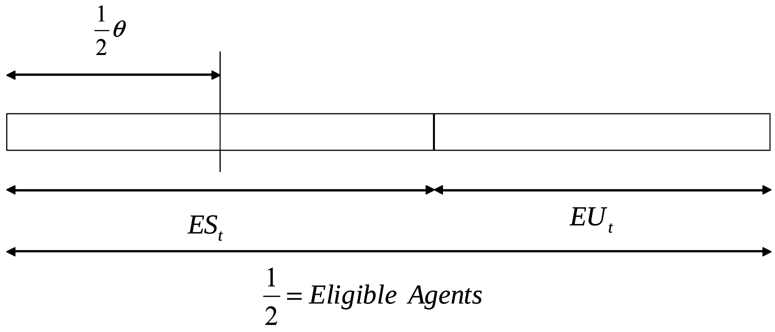


Figure 2: Selection in the first case

In such a case, the parameter θ is not important, since the relative proportion of agents who follow strategy A is the same for any θ. The reason for this is that the system selects a quantity $\theta\frac{1}{2}$ of agents randomly from among the ES_t. Since the proportion of agents who follow strategy A in ES_t is $\frac{ES_t^a}{ES_t}$, this will be the proportion of A-agents in the next level, if we simplify [11] that expression we have:

$$a_{t+1} = \frac{2P_1 a_t - P_1 a_t^2}{P_2 + 2P_1 a_t - (P_1 + P_2)a_t^2} \qquad if \qquad \theta\frac{1}{2} \leq ES_t \tag{1}$$

Second case. We now focus on the case $\theta\frac{1}{2} > ES_t$, in which the system selects all of the ES_t agents and randomly some of the EU_t agents. More specifically, $\left(\theta\frac{1}{2} - ES_t\right)$ agents are selected from EU. See Figure 3.

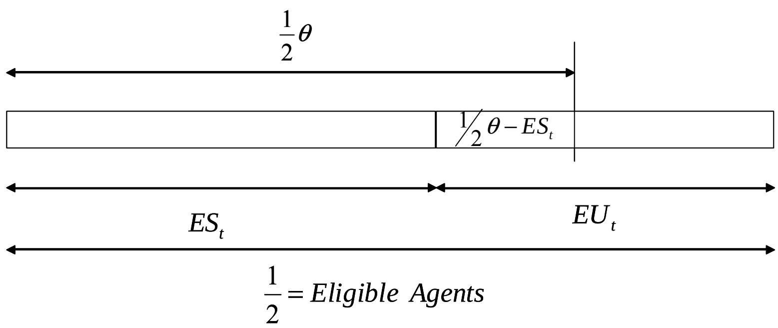


Figure 3: Selection in the second case

In this case, the A-agents selected by the system are all of the ES_t^a agents, plus the A-agents in the fraction of EU_t selected as well $\left(\text{i.e., } \theta\frac{1}{2} - ES_t\right)$. We already know ES_t^a

[11] Note $P3 = 1 - P1 - P2$ and $a_t = 1 - b_t$

but now we need to know the agents who are doing A within the fraction $\theta\frac{1}{2} - ES_t$ selected. The proportion of agents following strategy A in EU_t is $\frac{EU_t^a}{EU_t}$, thus, the number of A-agents who are selected is $\left(\theta\frac{1}{2} - ES_t\right)\frac{EU_t^a}{EU_t}$. Therefore, the total of A-agents selected is $ES_t^a + \left(\theta\frac{1}{2} - ES_t\right)\frac{EU_t^a}{EU_t}$, and the proportion of A-agents who are finally promoted is[12]:

$$\frac{1}{\theta\frac{1}{2}}\left(ES_t^a + \left(\theta\frac{1}{2} - ES_t\right)\frac{EU_t^a}{EU_t}\right)$$

From the previous expression, we obtain the following equation:

$$a_{t+1} = \frac{a_t}{\theta}(2P_1 - a_tP_1 + \tfrac{(1-P_1+(-1+a_t)P_2)(-2a_tP_1-P_2+a_t^2(P_1+P_2)+\theta)}{1-2a_tP_1-P_2+a_t^2(P_1+P_2)} \quad if \quad \theta\frac{1}{2} > ES_t \tag{2}$$

As such, the state at level $t+1$ is described by the following equation.

$$a_{t+1} = \begin{cases} f_1(a_t) & if \quad a_t2P_1 - a_t^2(P_1+P_2) + P_2 \geq \theta \\ f_2(a_t) & if \quad a_t2P_1 - a_t^2(P_1+P_2) + P_2 < \theta \end{cases} \tag{3}$$

Where $f_1(a_t)$ is the right-hand side of the equation 1 and $f_2(a_t)$ is the right-hand side of the equation 2.

Before studying the long-run behavior with the previous equation, we shall first consider a special case:

We assume that the number of agents selected (i.e. $\frac{1}{2}\theta$ agents) by $S_{[2,1,\theta]}$ is smaller than the number of eligible agents that have been successful (ES_t). Thus, we consider a sub-set of selection systems that are the most stringent ones among the family of selection systems $S_{[2,1,\theta]}$. Note that, since we consider θ a fix parameter, if ES_t changes with t, we might have, at any moment, that $\theta\frac{1}{2} > ES_t$. We could avoid this, however, as the number of ES_t agents has a minimum value[13] and it is $\min_{a_t}(ES_t) = \frac{1}{2}\min\{P_1, P_2\}$. Therefore, if $\theta \leq \hat{\theta} = \min\{P_1, P_2\}$, always $\theta\frac{1}{2} < ES_t$.

We get the following results for $\theta \leq \hat{\theta}$ (We name $S_{[2,1,\theta\leq\hat{\theta}]}$ the family of selection system under that assumption).

Proposition 1 *Let us assume* $\hat{\theta} = \min\{P_1, P_2\}$*, and consider the selection process* $S_{[2,1,\theta\leq\hat{\theta}]}$*, given by the equation* (1)*. The globally stable equilibrium is:*

$$a^*[S_{[2,1,\theta\leq\hat{\theta}]}] = \begin{cases} 0 & if \quad P_2 \geq 2P_1 \\ \frac{2P_1-P_2}{P_1+P_2} & if \quad 2P_1 > P_2 > \frac{1}{2}P_1 \\ 1 & if \quad P_2 \leq \frac{1}{2}P_1 \end{cases}$$

This result establishes that, if the degree of stringency is sufficiently high, i.e., the parameter θ is smaller than $\min\{P_1, P_2\}$, we can observe two characteristics in the behavior of the system in the long run: In the first place, the population profile does not depend, in the long run, on the degree of stringency. (The reason for this has already been explained). Secondly, if the difference between the expected success rates of the two different rules[14] is not high ($2P_1 > P_2 > \frac{1}{2}P_1$), in the long run, there are agents who continue following either rules. If that gap happens to be wide enough, however,

[12] Note that the total number of selected agents is $\theta\frac{1}{2}$. We therefore divide the total number of A-agents promoted by the total number of agents promoted.

[13] Since $ES_t = \frac{1}{2}a_t^2 P_1 + a_t b_t(P_1 + P_2) + \frac{1}{2}b_t^2 P_2 = a_tP_1 - \frac{1}{2}a_t^2(P1+P2) + \frac{P2}{2}$ is concave in a_t, the minimum is either in $a_t = 0$ or in $a_t = 1$. Note that $a_t \in [0, 1]$.

[14] The difference between the values P_1 and P_2.

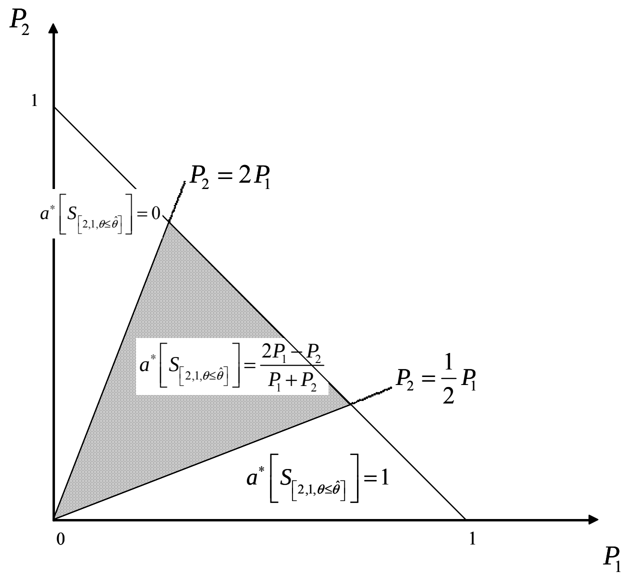


Figure 4: The three regions represent the three possible kind of outcomes in the long run.

in the long run, the entire population finally decides to follow the rule with the higher expected success rate (Figure 4).

Note that the higher the stringency degree, (i.e., the lower θ is), the more difficult it should be for the agent with bad outcomes to be promoted. Nevertheless, if the difference between the two rules is not too great, we find that even under the worst conditions for agents with bad outcomes, (i.e., with small θ), in the long run, there are still agents who keep following the rule with the lower expected success rate. Such agents manage to survive despite their lower expected success rates.

It seems reasonable to expect, therefore, that when the stringency degree is sufficiently low, (i.e., a high θ, since it affords the agents with bad outcomes a better chance of promotion), the agents with the lowest expected success rates also manage to survive in the long run. As the results we present below clearly show, however, this is surprisingly not the case. The more stringent the system is, (i.e., the smaller θ is), the better the agents with the worst success rates survive, provided that their "expected" performances or success rate gap ($|P_1 - P_2|$) is not too wide.

The following proposition gives us the globally stable equilibrium of the system under the selection process $S_{[n=2,k=1,\theta]}$. Without loss of generality, we consider the environment e_1 more common, $P_1 > P_2$, i.e., that the agents who follow Rule A have a greater success rate, in other words, a greater expected performance. The expression represented[15] by a^{**} depends on P_1, P_2 and θ.

[15] Expression a^{**} is the same as the root a_3 in the proof of Proposition 2, and it is the only feasible root of the equation $a_{t+1} - a_t = 0$, see Appendix.

$$a^{**} =$$

$$\frac{-\left((P_1^2-P_2^2)+(P_1-P_2)^2+(P_1+P_2)(1-\theta)\right)+\sqrt{\left((P_1^2-P_2^2)+(P_1-P_2)^2+(P_1+P_2)(1-\theta)\right)^2-4\left(P_2^2-P_1^2\right)(P_2-P_2^2+P_1(P_2+\theta-2))}}{2\left(P_2^2-P_1^2\right)}$$

Proposition 2 *Let* $P_1 > P_2$, $\bar{\theta}_1 = \frac{3P_1P_2}{P_1+P_2}$, $\bar{\theta}_2 = \frac{P_1^2+2P_2-P_1(1+P_2)}{P_2}$ *and*[16] *consider the selection process* $S_{[n=2,k=1,\theta]}$ *given by the equation*(3). *The globally stable equilibrium is:*

- *If* $P_2 > \frac{1}{2}P_1$ *then*

$$a^*[S_{[n=2,k=1,\theta]}] = \begin{cases} \frac{2P_1-P_2}{P_1+P_2} & when \quad \theta \leq \bar{\theta}_1 \\ a^{**} & when \quad \bar{\theta}_1 < \theta < \bar{\theta}_2 \\ 1 & when \quad \bar{\theta}_2 < \theta \leq 1 \end{cases}$$

- *If* $P_2 \leq \frac{1}{2}P_1$ *then*

$$a^*[S_{[n=2,k=1,\theta]}] = 1$$

Corollary 1 *If* $\theta \in [\bar{\theta}_1, \bar{\theta}_2]$ *then* $a^{**} \in [\frac{2P_1-P_2}{P_1+P_2}, 1]$. *Moreover,* a^{**} *is increasing in* θ.

The proposition outlined above establishes that for great differences between the success rate of the two rules followed by the agents ($P_2 \leq \frac{1}{2}P_1$) in the long run, all of the agents will follow the rule with the best success rate[17] (A). Thus, the value of θ does not really matter, as we get the same result (in the long run) for any θ. In other words, the degree of stringency does not matter in the long run. As such, the relative advantage of one rule is too great to be counteracted by any degree of stringency in the promotion system.

On the other hand, if that gap between the two success rates P_1 and P_2 is not too high ($P_2 > \frac{1}{2}P_1$), in the long run, the population profile depends on the value of the parameter θ, i.e., the degree of stringency in the system:

- If θ is smaller than a certain threshold ($\bar{\theta}_1$), the agents with the worst success rates (B) survive in the long run. Moreover, the proportion of agents following any given strategy does not change in the long run with θ. Notice that for agents who follow the strategy with the lower success rate, this scenario is the best one. They survive in greater proportions than in any other specification of the parameter θ.

- If θ is located between the two thresholds, ($\bar{\theta}_1 < \theta < \bar{\theta}_2$), the long-run behavior depends on the parameter θ. If θ increases (i.e., if the stringency degree decreases), the proportion of agents who follow the strategy with the worst success rate decreases, and if it is equal to or greater than the second threshold, they completely disappear in the long run.

[16] If $P_2 > \frac{1}{2}P_1 \Rightarrow \left[P_1 < \frac{3P_1P_2}{P_1+P_2} < \frac{P_1^2+2P_2-P_1(1+P_2)}{P_2}\right]$ (see proof of Proposition 2 in the Appendix). Moreover, obviously $\hat{\theta} < \bar{\theta}_1$

[17] When we say "success rate" of a rule, we refer to the "expected success rate" of following this rule. It holds in all sections.

Therefore, the proposition and corollaries outlined above establish, surprisingly enough, that the more stringent the system is (i.e., the smaller θ is), the better the chances the agents with the worst success rates have of surviving, (provided that the gap between their "expected" performances, or success rates, is relatively low). This occurs in spite of the fact that the probability of agents with bad outcomes being promoted increases provided that the stringency degree decreases (θ increases). In other words, if more (fewer) people are promoted, it is easier (more difficult) for agents with bad outcomes to be promoted, and, obviously, agents who follow the rule with the worst success rate obtain bad outcomes more often than those who follow the one with the best success rate. In the long run, however, the greater the stringency is, the higher the number of agents who follow the rule with the worst success rate.

To give a clearer picture of just what is happening, we shall first try to simplify the matter a bit. We continue considering $P_1 > P_2$, and as such, we associate agents who follow the $A(B)$ rule (labelled $A(B)$ -agents) with agents with high (low) success rates. We shall now focus on their chances of promotion. Note that, the population of $A(B)$-agents increases if the probability of promotion is greater for $A(B)$-agents than it is for $B(A)$ agents. If both probabilities are equal, the population ratios do not change and the system is therefore in equilibrium.

Provided $\theta = 1$, the only globally stable equilibrium is $a = 1$, when all of the agents are following Rule A. If θ decreases, however, this equilibrium becomes unstable and the system converges to a stable globally mixed equilibrium in which there are agents following both rules.

We now focus on the system in which $a_t \simeq 1$, in which case, if θ is relatively small ($\theta < \hat{\theta} = P_2$), the probability of an A-agent being promoted is $\frac{1}{2}P_1$. This is because an A-agent is matched with another A-agent with a probability close to one. There is therefore a probability of $\frac{1}{2}$ that an A-agent will pass the local selection, (i.e., of being an eligible agent). Since θ is relatively small, the A-agent must necessarily be a successful eligible agent for him to be promoted to the next level, which occurs with a probability of P_1. On the other hand, a B-agent also has a probability close to one of being matched against an A-agent. Thus, there is a probability of P_2 that he will pass the local selection. For B-agents, the only possibility of passing the local selection is by responding correctly to his environment. Therefore, if $P_2 > \frac{1}{2}P$, then $a = 1$ is unstable.

In conclusion, if $a_t \simeq 1$, the local selection is harder for the B-agents than for the A-agents. As a result almost all of the B-agents who pass the local selection are successful, which, in fact, is their only possibility of passing. If, therefore, the level of global stringency increases, (i.e., if θ decreases), the ones who are negatively affected are the A-agents, and more specifically, the unsuccessful eligible A-agents.

3.2 The promotion system $S[n = 2, k = 2, \theta]$.

Since we have considered (in the previous section) the size of the local group to be equal to 2 ($n = 2$), we can only consider two degrees of local promotion; $\frac{k}{n} = \frac{1}{2}$ and $\frac{k}{n} = 1$ and, therefore, two degrees of local stringency $\frac{1}{2}$ and 0. As the former has already been studied in the previous section, we shall now study the latter case, for which, we should emphasize, there is no local selection. The entire population of a level who are matched for promotion will be eligible agents. We follow the same method as before to obtain the dynamic equation.

$$
\begin{aligned}
ES_t &= a_t\, P_1 + b_t P_2 \\
ES_t^a &= a_t\, P_1 \\
EU_t &= a_t\,(P_2 + P_3) + b_t(P_1 + P_3) \\
EU_t^a &= a_t\,(P_2 + P_3)
\end{aligned}
$$

Note that all the agents are eligible and as before, so that a proportion (P_1) of the A-agents ($a_t P_1$), as well as a proportion (P_2) of the B-agents ($b_t P_2$) will always be successful. If we consider other promotion systems with greater values of k or n, the previous sentence is quite true, provided that $k = n$. Therefore, the previous equations will be the same for any value of k or n, provided that $k = n$, (i.e., that all the promotion systems in which $k = n$ will have the same dynamics).

The following equation describes the state at level $t + 1$:

$$
a_{t+1} = \begin{cases} \frac{a_t P_1}{a_t(P_1 - P_2) + P_2} & if \quad a_t(P_1 - P_2) + P_2 \geq \theta \\ \frac{a_t(P_1 - P_2 + (1-P_1)\theta) + a_t^2(P_1 - P_2)}{(1 - a_t(P_1 - P_2) - P_2)\theta} & if \quad a_t(P_1 - P_2) + P_2 < \theta \end{cases} \tag{4}
$$

The next proposition shows us that if there are not local selection, i.e.,the local stringency degree is equal to 0, the agents who follow the rule with the higher success rate will finally be the only survivors.

Proposition 3 *Let $P_1 > P_2$ and consider the selection process $S_{[2,2,\theta]}$ given by the equation*(4)*.*

The globally stable equilibrium is:

$$
a^*[S_{[2,2,\theta]}] = 1
$$

It would seem, therefore, that a change in the local stringency degree has a similar effect to a change in the global stringency degree. If the local stringency changes from $\frac{1}{2}$ (it was studied in the previous section) to 0, (i.e., if there is no local selection), the only survivors, in the long run, are the A-agents.

If we really wish to understand the effect of the degree of local stringency, however, we must study a more general case, with larger local groups, as it allows us to consider different degrees of local stringency.

4 Groups of n agents with $k = 1$, $S[n, k = 1, \theta]$.

In this section, we study the case of $S[n, k = 1, \theta]$. We assume $k = 1$ and consider n as a parameter. As such, we analyze the effect of an increase in the size of local groups on the long-run population profile. The agents now interact in groups of n agents. In the previous section we have studied a particular case of this scenario. If we consider different group sizes, we can also consider different local stringency degrees. We have already shown that the degree of local stringency could be measured by the quotient $\frac{k}{n}$, in fact, it is equal to $1 - \frac{k}{n}$. By changing the parameter n, therefore, we can consider promotion degrees between[18] 0 and $\dfrac{1}{2}$.

[18] Since we consider in this section $k = 1$, the degree of local promotion will be $\dfrac{k}{n} \in \left(0, \dfrac{1}{2}\right]$, more specifically $\dfrac{k}{n} \in \left\{\dfrac{1}{n}, \forall n \in N\right\}$, thus the degree of local stringency will be in the interval $[\dfrac{1}{2}, 1)$.

Before obtaining the dynamic equations, we need to know how many groups there are. Note that we have a_t agents following Rule A and b_t following Rule B at level t, with $a_t + b_t = 1$. Since they are the agents who must compete to be promoted, they are randomly matched in groups of n. Thus, the probability that there is an A-agent in one of these groups is equal to a_t. Therefore, the probability of there being a group with a number x of A-agents (and with a number $k - n$ of B-agents) is[19]:

$$\binom{n}{x} a_t^x (1 - a_t)^{n-x}$$

Since there is a population of agents that is equal to one, there are $\frac{1}{n}$ groups of n agents. Therefore, there are $\frac{1}{n}\binom{n}{x} a_t^x (1 - a_t)^{n-x}$ groups with x agents following Rule A.

As such, we can now obtain the following expressions:

$$\begin{aligned}
ES_t &= \sum_{x=1}^{n} \frac{1}{n}\binom{n}{x} a_t^x (1 - a_t)^{n-x} P_1 + \sum_{x=0}^{n-1} \frac{1}{n}\binom{n}{x} a_t^x (1 - a_t)^{n-x} P_2 \\
ES_t^a &= \sum_{x=1}^{n} \frac{1}{n}\binom{n}{x} a_t^x (1 - a_t)^{n-x} P_1 \\
EU_t &= \frac{1}{n} - ES_t \\
EU_t^a &= a_t^n (P_2 + P_3) + \sum_{x=1}^{n-1} \frac{1}{n}\binom{n}{x} a_t^x (1 - a_t)^{n-x} \frac{x}{n} P_3
\end{aligned}$$

To obtain the number of agents who have passed the local selection and, in addition, have been successful after the matching up in their groups (i.e., ES_t), we must consider the following facts: First, since one agent from each group passes the local selection ($k = 1$), the number of agents who are finally selected is equal to the number of groups there are. Secondly, all of the different kinds of groups have A-agents except one, which is the kind of group in which all the agents are B-agents (in which case, $x = 0$). Therefore, the proportion[20] P_1 of the groups with A-agents will be equal to the number of agents who pass the local selection successfully and, moreover, follow Rule A. We work along the same lines to obtain the B-agents who pass the local selection successfully.

Obviously, the number of agents who pass the local selection unsuccessfully[21] (EU_t) will be equal to the number of agents who pass the selection $\left(\frac{1}{n}\right)$ minus the agents who pass it successfully (ES_t). Finally, the number of A-agents who pass the local selection unsuccessfully (EU_t^a) will be equal to: (1) the number exclusively A-agent groups (a_t^n) under either Environment e_2 or Environment e_3 (which occurs with a probability of $(P_2 + P_3)$), and (2) we must add up the groups with both kinds of agents under Environment e_3, in which case, all of the agents have the same probability of being promoted, all agents in those group gave the bad respond to the environment. Therefore, the probability that the agent selected would be an A-agent is equal to $\frac{x}{n}$.

Analogously to the previous sections, for $n = 2$, the proportion of A-agents in Level $t + 1$ must be:

[19] The x is distributed as a binomial distribution, $x \backsim B(n, a_t)$.

[20] A proportion P_1 of the groups are under the environment e_1, in which case, Rule A is the right response to the environment.

[21] They gave the bad respond to their environment.

$$a_{t+1} = \begin{cases} \frac{ES_t^a}{ES_t} & if \quad ES_t \geq \frac{\theta}{n} \\ \frac{1}{\theta \frac{1}{n}} \left(ES_t^a + \left(\theta \frac{1}{n} - ES_t \right) \frac{EU_t^a}{EU_t} \right) & if \quad ES_t < \frac{\theta}{n} \end{cases}$$

Therefore, if we replace those expressions presented above, we obtain the dynamic equation for the promotion system $S[n, k = 1, \theta]$:

$$a_{t+1} = \begin{cases} \frac{(1-(1-a_t)^n)P_1}{(1-(1-a_t)^n)P_1+(1-a_t^n)P_2} & if \quad (1-(1-a_t)^n)P_1+(1-a_t^n)P_2 \geq \theta \\ \frac{1}{\theta}\left[(1-(1-a_t)^n)P_1+(a_t^n P_2+a_t(1-P_1-P_2))\frac{\theta-(1-(1-a_t)^n)P_1+(1-a_t^n)P_2}{1-(1-(1-a_t)^n)P_1+(1-a_t^n)P_2}\right] & if \quad (1-(1-a_t)^n)P_1+(1-a_t^n)P_2 < \theta \end{cases} \tag{5}$$

To understand what happens in the long run, we consider two extreme cases, one in which there is no global selection, (i.e., $\theta = 1$), and another in which there is the strongest global selection[22] (i.e.,$\theta \leq \hat{\theta} = \min\{P_1, P_2\}$).

4.1 The promotion system $S[n, k = 1, \theta = 1]$

In such a case, equation (5) will be:

$$a_{t+1} = (1 - (1 - a_t)^n)P_1 + a_t^n P_2 + a_t(1 - P_1 - P_2) \tag{6}$$

Here we find two possible kinds of long-run behavior, either homogeneous, with the entire population following the same rule, or heterogeneous, with both rules persisting in the long run. It is not possible, however, to find an explicit expression for the heterogeneous equilibrium, which we denote by a^*_{M1}. Where a^*_{M1} is a root of the equation:

$$a_{t+1} - a_t = (1 - (1 - a_t)^n)P_1 + a_t^n P_2 - a_t(P_1 + P_2) = 0 \tag{7}$$

This root belongs to the open interval $(0, 1)$, and it exists and is unique provided that $\frac{1}{(n-1)}P_1 > P_2 > (n - 1)P$. (See appendix).

Proposition 4 *Consider the selection process $S_{[n,1,\theta=1]}$ given by the equation*(6)*. The globally stable equilibrium is:*

$$a^*[S_{[n,1,\theta=1]}] = \begin{cases} 0 & if \quad P_2 \geq (n-1)P_1 \\ a^*_{M1} & if \quad \frac{1}{(n-1)}P_1 < P_2 < (n-1)P_1 \\ 1 & if \quad P_2 \leq \frac{1}{(n-1)}P_1 \end{cases}$$

Corollary 2 *Let $P_1 > P_2$. In this case $a^*[S_{[n,1,\theta=1]}]$ is decreasing in n. Thus, $a^*[S_{[n,1,\theta=1]}] \geq a^*[S_{[n+1,1,\theta=1]}]$ and the inequality is strict, provided that $\frac{1}{(n-1)}P_1 < P_2$.*

Let us suppose that $P_1 > P_2$. In such a case, the greater the gap between the success rates ($|P_1 - P_2|$) is, the greater the advantage obtained from following Rule A is. As we can see from the previous proposition, if the gap between the success rates is sufficiently high, ($P_2 \leq \frac{1}{(n-1)}P_1$), the entire population follows Rule A in the long run.

On the other hand, in the local competition, an agent has to compete with $n - 1$ others and if that number increases, (i.e., if n increases), the increased competition in

[22] As in the previous section, with $n = 2$, the minimum value of the expresion $ES_t = \frac{1}{n}((1 - (1 - a_t)^n)P_1 + (1 - a_t^n)P_2)$ with $a_t \in [0, 1]$ is equal to $\frac{1}{2}\min\{P_1, P_2\}$.

local selection makes it more difficult for an individual agent to pass it. An increase of n produces an increase in the local stringency degree of the selection process, i.e., $\frac{k}{n}$ decreases.

The previous proposition demonstrates that no matter how large the difference between the success rates ($|P_1 - P_2|$) of the two different rules is, if n increases enough, the rule with the lower success rate survives in the long run. (See Figure (5)).

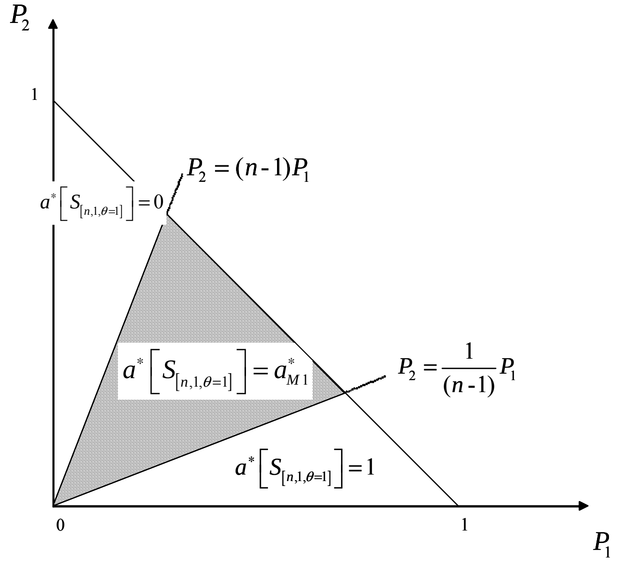


Figure 5: The three regions represent the three possible qualitative outcomes in the long run. Note that if n increases, the shaded region, where both kinds of agents survive in the long run, also increases.

The previous corollary establishes that, if n increases, the proportion of agents who follow the rule with the lower success rate increases in the long run.

In short, no matter how small the success rate of a rule is, if the local stringency degree is high enough, this rule survives in the long run. Moreover, the greater the local stringency degree is, the larger the proportion of agents who follow the rule with the lower success rate will be. This might well seem paradoxical, but the explanation for it is outlined in the next section.

To conclude this section, note that if n goes to infinity, equation (6) goes to the equation $a_{t+1} = P_1 + a_t(1 - P_1 - P_2)$, and therefore, $a^*[S_{[n,1,\theta=1]}]$ goes to $\frac{P_1}{P_1+P_2}$. Thus, if $P_1 > P_2$ then $a^*[S_{[n,1,\theta=1]}] \in (\frac{P_1}{P_1+P_2}, 1]$. In the Figure 6 we can see the dynamic function(6) with different values of n.

4.2 The promotion system $S[n, k = 1, \theta \leq \hat{\theta} = \min\{P_1, P_2\}]$.

In such a case, the number of agents who are selected in the global selection $(\frac{1}{n}\theta)$ is always smaller than the number of agents who pass the local selection (ES_t). All of the agents who were promoted, therefore, were successful in their environments. Thus, we can rewrite the equation (5) in the following way.

$$a_{t+1} = \frac{(1 - (1 - a_t)^n)P_1}{(1 - (1 - a_t)^n)P_1 + (1 - a_t^n)P_2} \tag{8}$$

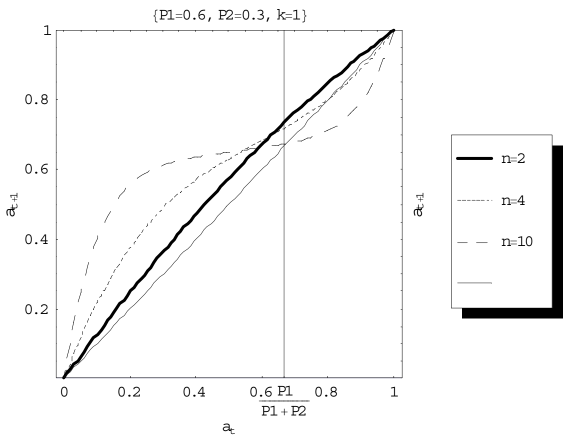


Figure 6: The graph represents three different dynamic functions $a_{t+1} = f(a_t)$ for the selection processes $S_{[n,1,\theta=1]}$ with $P_1 = 0.6$ and $P_2 = 0.3$. We consider three functions with three different sizes of groups, $n = 2$, $n = 4$ and $n = 10$.

As in the previous section, we find two possible kinds of long-run behavior, either homogeneous, with the entire population following the same rule, or heterogeneous, with both rules persisting in the long run. It is not possible to find an explicit expression for the heterogeneous equilibrium, which we denote by a^*_{M2}. Where a^*_{M2} is a root of the equation:

$$a_{t+1} - a_t = \frac{(1-(1-a_t)^n)P_1}{(1-(1-a_t)^n)P_1 + (1-a_t^n)P_2} - a_t = 0 \tag{9}$$

This root belongs to the open interval $(0, 1)$ and this root only exists and is unique provided that $\frac{1}{n}P_1 > P_2 > nP_1$.

We obtain the following result:

Proposition 5 *Consider the selection process $S_{[n,1,\theta\leq\hat{\theta}=\min\{P_1,P_2\}]}$ given by the equation* (8)*. The globally stable equilibrium is:*

$$a^*[S_{[n,1,\theta\leq\hat{\theta}=\min\{P_1,P_2\}]}] = \begin{cases} 0 & \mathit{if} \quad P_2 \geq nP_1 \\ a^*_{M2} & \mathit{if} \quad \frac{1}{n}P_1 < P_2 < nP_1 \\ 1 & \mathit{if} \quad P_2 \leq \frac{1}{n}P_1 \end{cases}$$

Note that, the previous equation(9) is equivalent to $\frac{(1-(1-a_t)^{n+1})P_1 + a_t^{n+1}P_2 - a(P_1+P_2)}{(1-(1-a_t)^n)P_1+(1-a_t^n)P_2} = 0$ and it has the same roots as the equation $(1-(1-a_t)^{n+1})P_1 + a_t^{n+1}P_2 - a(P_1+P_2) = 0$, and there is only one difference between this last equation and equation (7), which is that $n+1$ appears instead of n. We can therefore state that:

$$a^*[S_{[n+1,1,\theta=1]}] = a^*[S_{[n,1,\theta\leq\hat{\theta}]}]$$

Moreover, since the equilibrium is decreasing in n (provided $P_1 > P_2$):

$$a^*[S_{[n,1,\theta=1]}] \geq a^*[S_{[n,1,\theta\leq\hat{\theta}]}]$$

Thus, we know that if θ changes from 1 to $\hat{\theta}$, the global stringency increases. Likewise, if n changes to $n+1$, the local stringency degree increases. As we can see, therefore, in the long run the effect is the same in both cases. If we increase the size of the groups by one individual, the effect, in the long run, is the same as imposing the highest degree of global stringency (from $\theta = 1$ to $\theta \leq \hat{\theta}$).

On the other hand, note that if n goes to infinity, the equation (8) goes to the equation $a_{t+1} = \frac{P_1}{P_1+P_2}$, and therefore, $a^*[S_{[n,1,\theta\leq\hat{\theta}]}]$ goes to $\frac{P_1}{P_1+P_2}$. Thus, if $P_1 > P_2$ then $a^*[S_{[n,1,\theta\leq\hat{\theta}]}] \in (\frac{P_1}{P_1+P_2}, 1]$

In conclusion, the two scenarios outlined above are both extreme cases (in θ) of the family of selection processes $S[n, 1, \theta]$, i.e., there is no global selection at one extreme ($\theta = 1$) and at the other extreme, the global selection is the hardest ($\theta \leq \hat{\theta}$). Therefore, the behavior of the systems with any θ must be somewhere between the two extremes. As we have seen, therefore, the effect of an increase in the degree of stringency in selection processes, whether local or global, is an increase in the proportion of agents who follow the rule with the lower success rate in the long run, provided that the difference between the success rates between the two rules is not too high.

So far, however, we have only considered degrees of stringency that are greater than $\frac{1}{2}$, i.e., degrees of promotion smaller than $\frac{1}{2}$. Note also that the agents who are promoted to the next level are from a population that has been normalized to 1. Therefore, the stringency degree[23] is $\frac{1}{n}\theta$, the product between the local and global degrees of stringency. We now wish to know what happens when the stringency degree is smaller than $\frac{1}{2}$ i.e., the proportion of agents promoted are greater than $\frac{1}{2}$. To find the answer to this question we are going to study, (in the following section), the more general family of promotion systems $S[n, k, \theta]$, in which case, the promotion degree is equal to $\dfrac{k}{n}\theta$.

5 The general case $S[n, k, \theta]$.

We now consider the more general promotion system. The objective is to know what happens with low levels of stringency. As we have already mentioned, in this case, the agents are randomly matched in groups of n agents. The best k agents in each group are selected after the realization of the idiosyncratic environment. These are the eligible agents, i.e., those who have passed the local selection. Afterwards, a proportion θ of all the eligible agents are selected for promotion, which is the global selection. Therefore, a proportion $\frac{k}{n}\theta$ of all the agents who compete for promotion are eventually promoted to the next level.

In this section, we simplify matters by considering that:

$$P_3 = 0 \text{ (i.e. } P_1 + P_2 = 1) \text{ and } P_1 > P_2$$

Therefore, Rule A has a greater success rate than Rule B has, and we can obtain the following expressions in the same way as we did in the previous section.

[23] It should be remembered that, for us, the stringency degree of a promotion system is parameterized by the proportion of agents promoted (one minus this proportion). As the system first promotes the agents who have had the best outcomes, the more stringent the system is, the more difficult promotion will be for agents with bad outcomes.

$$
\begin{aligned}
ES_t &= \sum_{x=1}^{k-1} x\frac{1}{n}\binom{n}{x}a_t^x(1-a_t)^{n-x}P_1 + k\sum_{x=k}^{n}\frac{1}{n}\binom{n}{x}a_t^x(1-a_t)^{n-x}P_1 + \\
&\quad k\sum_{x=0}^{n-k}\frac{1}{n}\binom{n}{x}a_t^x(1-a_t)^{n-x}P_2 + \sum_{x=n-k+1}^{n-1}(n-x)\frac{1}{n}\binom{n}{x}a_t^x(1-a_t)^{n-x}P_2 \\
ES_t^a &= \sum_{x=1}^{k-1} x\frac{1}{n}\binom{n}{x}a_t^x(1-a_t)^{n-x}P_1 + k\sum_{x=k}^{n}\frac{1}{n}\binom{n}{x}a_t^x(1-a_t)^{n-x}P_1 \\
EU_t &= \frac{k}{n} - ES_t \\
EU_t^a &= \sum_{x=n-k+1}^{n}(k-(n-x))\frac{1}{n}\binom{n}{x}a_t^x(1-a_t)^{n-x}P_2
\end{aligned}
$$

As we have already shown, there is a proportion $\binom{n}{x}a_t^x(1-a_t)^{n-x}$ of groups with x agents who follow Rule A (A-agents) and there are $\frac{1}{n}$ groups in a level. After the realization of the environment, these groups can now be divided into two different types: groups under the environment e_1, (a proportion P_1 of all the groups), and groups under the environment e_2, (a proportion P_2 of all the groups).

If we wish to know ES_t^a, i.e., the number of A-agents who have passed the local selection, and who, furthermore, have been successful in their groups, we must consider only the groups under environment e_1. Obviously, Rule A is the successful rule for these groups alone. If the group has a number $x < k$ of A-agents, therefore, a number x of A -agents will be chosen, and if $x \geq k$, then a number k of A-agents are selected from the groups under environment e_1. Thus, such agents are the successful eligible agents who follow Rule A (ES_t^a) and who have passed the local selection.

We can work in similar way to obtain ES_t $(= ES_t^a + ES_t^b)$, and EU_t^a. To obtain the dynamic equation $a_{t+1} = f(a_t)$ we merely have to replace the previous expressions in the following equation:

$$
a_{t+1} = \begin{cases} \frac{ES_t^a}{ES_t} & if \quad ES_t \geq \frac{k}{n}\theta \\ \frac{1}{\theta\frac{k}{n}}\left(ES_t^a + \left(\theta\frac{k}{n} - ES_t\right)\frac{EU_t^a}{EU_t}\right) & if \quad ES_t < \frac{k}{n}\theta \end{cases} \tag{10}
$$

The expression we obtain after the replacement, however, is far too large and complex and, furthermore, we did not consider it interesting enough to include it in this paper. Moreover, the complexity of the equation makes it impossible for us to work in the same way as we did in the previous section.

We use numerical calculus to discover the behavior of the system in the long run. We find that, in the long run, there are three different kinds of behavior, which are shown in Figure 7.

Obviously, $a = 0$ and $a = 1$ are always equilibria or stationary points of the system. Moreover, there are some cases in which one other (inner) equilibrium arises[24] , i.e., it belongs to the open interval $(0, 1)$. The existence of this inner equilibrium and the stability of all equilibria depend on the parameters $\{P_1, P_2, n, k, \theta\}$.

As can be seen in the Figure 7, there are three possible types of long-run behavior: (a) an inner equilibrium exists and it is globally stable, (b) there is no inner equilibrium

[24] We find by numerical calculus that the equation $f(a_t) - a_t = 0$ has *at most* one solution in the open interval $a_t \in (0, 1)$.

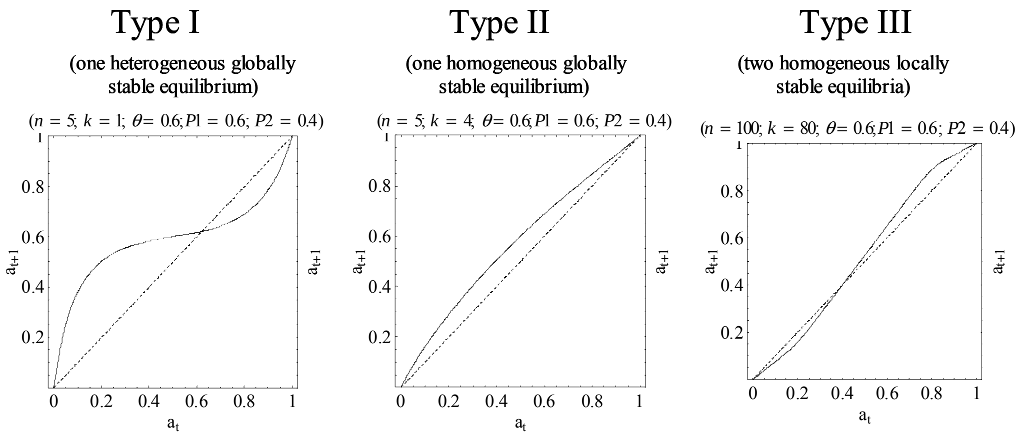


Figure 7: We represent the dynamical function $a_{t+1} = f(a_t)$ for three different parameter specifications.

and $a = 1$ is globally stable, and (c) an inner equilibrium exists which is unstable. In this latter case, there are two basins of attraction, and $a = 0$ and $a = 1$ are locally stable. We find the first sort of behavior only when the stringency degree is high. If it increases, however, we then find the second sort of behavior, and the two basins of attraction appear only when the stringency degree is low enough.

We now wish to discover precisely when the system has any of the three different kinds of behavior. To do so, we analyze the stability of the equilibria $a = 0$ and $a = 1$. The stability of these equilibria determines the existence and stability of the inner equilibrium since the system can only have one of the three different kinds of behavior that we have found by numerical calculus.

Note that if the system is in a period t (or level t), the proportion of A-agents increases or decreases in $t+1$ depending on the probabilities for promotion of the agents in t. In others words, if the promotion probability of an A-agent is greater than that of an B-agent in period (or level) t, then the proportion of A-agents increases in the next period (or level). Obviously, in such a case, the proportion of B-agents decreases.

If we wish to know when Equilibria ($a = 0$ and $a = 1$) are either stable or unstable, therefore, we can study the promotion probabilities, outlined above, when the system are close to these equilibria. For example, if we prove that the probability of promotion for an A-agent is greater (smaller) than it is for a B-agent, provided that the system is indefinitely very close to $a = 0$, then we can state that these equilibria are unstable (stable). We can do the same thing with $a = 1$. Thus, we obtain the following result, which explains when Equilibria $a = 0$ and $a = 1$ are either (locally) stable or unstable. (See appendix).

Proposition 6 *Let $P_1 > P_2$, $P_3 = 0$ and consider the selection process $S_{[n,k,\theta]}$ given by the equation*(10) *.*

If $\theta < P_1$ and,

$\frac{k}{n} < \frac{P_2}{P_1}$ *then $a = 0$ and $a = 1$ are unstable.*

$\frac{k}{n} \geq \frac{P_2}{P_1}$ *then $a = 0$ is unstable and $a = 1$ is stable.*

If $\theta \geq P_1$ and,

$\frac{k}{n}\theta < P_2$ *then* $a = 0$ *and* $a = 1$ *are unstable.*

$P_2 \leq \theta \leq P_1$ *then* $a = 0$ *is unstable and* $a = 1$ *is stable.*

$\frac{k}{n}\theta > P_1$ *then* $a = 0$ *and* $a = 1$ *are stable.*

These results show that if the global promotion degree θ is smaller than the success rate of Rule A , i.e., the e_1 probability(P_1), then the qualitative behavior does not depend on θ in the long run. It only depends on the local promotion degree $\frac{k}{n}$, and, if it is smaller than $\frac{P_2}{P_1}$, then both equilibria are unstable. Thus, the inner equilibrium[25] a_M exists and is globally stable, $a^*[S_{[n,k,\theta]}] = a_M$, long run behavior $Type\ II$.

On the other hand, if $\theta \geq P_1$, in the long run, the qualitative behavior depends on the absolute promotion degree $\frac{k}{n}\theta$, i.e., both local and global. In this case, if the absolute promotion degree is smaller than P_2 then the long-run behavior is of $Type\ I$. If it is between P_1 and P_2, it is of $Type\ II$, and if it is greater than P_1, there are two stable equilibria and we find a long-run behavior of $Type\ III$.

We can resume that in the Figure[26]8. There are two cases: first [27] $P_1 > \frac{P_2}{P_1}$. Secondly, we have $P_1 < \frac{P_2}{P_1}$.

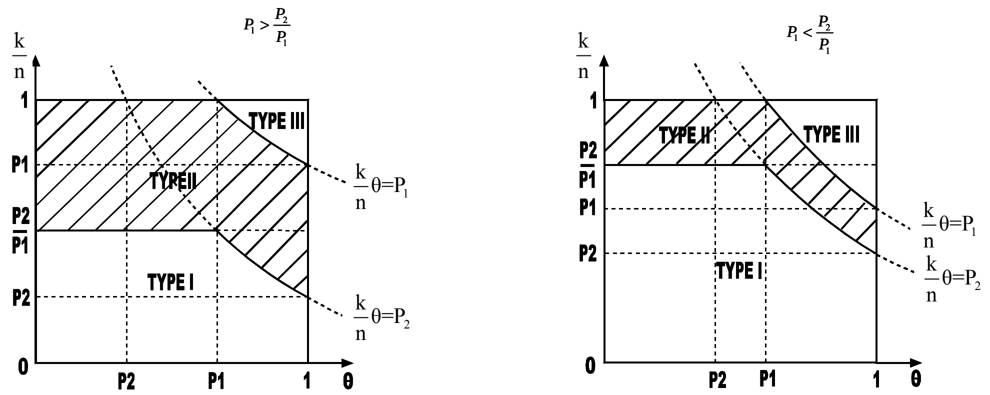


Figure 8: The vertical axes of these graphs represent the local promotion degree ($\frac{k}{n}$), the horizontal axes represent the global promotion degree (θ). The right graph in the case of $P_1 < \frac{P_2}{P_1}$ and the left graph in the case of $P_1 > \frac{P_2}{P_1}$. We have three regions which represents the three types of behaviours. TYPE I: $a = 0$ and $a = 1$ are unstable, the inner equilibrium a_M exists and is globally stable. TYPE II: $a = 0$ is unstable and $a = 1$ is globally stable. TYPE III: $a = 0$ and $a = 1$ are locally stable, a_M exists and is unstable.

In conclusion, given P_1 and P_2, i.e., given the relative advantage of following a given rule, an increase in the stringency degree (i.e., $\frac{k}{n}\theta$ decreases) has a negative effect on the homogeneity of the population[28] in the long run. As can be seen in Figure 8, if the

[25]Note that, a_M can depend on θ in the long run, although, in such a case $\theta < P_1$, the existence of a_M does not depend on θ.

[26]As we have already seen in Section 3.2, if $\frac{k}{n} = 1$ then $a^*[S_{[n,k,\theta]}] = 1$, provided that $P_1 > P_2$. Therefore, the line $\frac{k}{n} = 1$ is in region of Type II.

[27]We have that $P_1 > \frac{P_2}{P_1} \Leftrightarrow P_1 > \frac{(1-P_1)}{P_1} \Leftrightarrow P_1^2 + P_1 - 1 > 0$, which occurs provided that $P_1 > \frac{\sqrt{5}-1}{2} \simeq 0.618$

[28]If all the agents follow the same rule in the long-run equilibrium, it is homogeneous. If the two rules happen to be in equilibrium, however, the equilibrium is heterogeneous. The more equal the proportions of agents who follow the two different rules in equilibrium are, the more heterogeneous the equilibrium will be.

stringency degree of the promotion system is low enough, the homogeneous equilibria ($a = 0$ and $a = 1$) are both stable, *Type III*. In such a case, there are two basins of attraction. As the stringency degree increases, the homogeneous equilibria become unstable, beginning with the weaker homogeneous equilibrium[29], *Type II*. Finally, if the stringency degree of the system is high enough, both homogeneous equilibria are unstable and there is an inner or heterogeneous equilibrium that is globally stable, *Type I*. Moreover, the higher the stringency degree is, the more heterogeneous the inner equilibrium will be, which means, *no matter how low the success rate of a rule is, if the stringency degree is high enough, the rule survives in the long run.*

In summary then, the more stringent the system is, the more heterogeneous the long-run behavior is. Thus, if the system is stringent enough, the agents who follow the less-successful rule manage to survive in the long run. If the stringency of the system is maintained between two given thresholds, the agents who follow the more successful rule are the only survivors. Finally, if the stringency of the system is low enough, both homogeneous equilibria are stable and the dynamics depend on the initial conditions.

To understand why this happens, we must first observe that the dynamics of the system depends on the agents' probabilities of promotion. The proportion of the kind of agents who have the greatest probability of promotion in a period (or level) t will increase in $t + 1$. Let us now consider a certain type of agent who faces one of the two following extreme scenarios:

- If a particular type of agent is scarce, (let us say close to extinction), the fact that the degree of stringency in the system increases does not affect the probability of promotion of this type of agent. Note that, in general, there would only be one agent of this type in a group, and such an agent would only be promoted if he was successful (and, therefore, the rest of agents in his group would have to be unsuccessful). Thus, if this scarce agents is successful (gives the right respond to the environment), he will promote almost for sure.

- However, when a given type of agent is quite abundant, (i.e., close to homogeneity), the fact that the degree of stringency in the system increases does affect a lot the probability of promotion of such type of agents. Note that, in general, there would only agents of this type in the different groups. Thus, if all (or close to all) members of a group are of the same type and, therefore, they give the same response under the same environment, then it does not matter at all if one is successful or not. The reason for this is that all of his competitors in his own group are also successful (or unsuccessful) as he is. The individual's probability of promotion, therefore, depends mainly on how many people are promoted.

An increase in stringency, therefore, tends to punish the more common type of agent, while it does not affect the relatively scarce type. Thus, beyond a certain point, it can cause the homogenous equilibrium (made up of the more successful agents) to lose its stability.

On the other hand, if the degree of stringency is low enough, even the homogeneous equilibrium of the less-successful agents can be stable. This explains why both homogeneous equilibria are locally stable in this particular case. Consequently, the dynamics depend on the initial conditions, and as such, the considerations pertaining to relative numbers are so important that the long-run behavior depends on the initial conditions.

[29] If $P_1 > P_2$, then $a = 0$ is the weaker homogeneous equilibrium, the B-agents are less successful on average than the A-agents.

The degree of stringency permits a control over the diversity of a system, and can either prevent or stimulate the homogeneity of this system.

The less stringent a system is, the easier it is for it to be dominated by one type of agent and for it to achieve homogeneity.

The more stringent a system is, the more important an agent's success or failure in the promotion becomes, and as such, the less important the effect of the relative proportions of the different types of agents is. As we have already seen, if the stringency degree is low enough, the long-run behavior depends exclusively on the initial conditions.

The **fundamental assumption** is that the successes (or failures) of agents who follow the same rule correlate perfectly among themselves. If two agents are under the same environment and are following the same rule and one of them is successful, then the other one will be successful as well, with a probability of one.

6 The Optimum.

In the previous sections we have studied the long-run behavior of a hierarchical social system with different promotion mechanism. We now wish to know something about the efficiency of the selection process to be able to decide on the optimal sort of composition for a population. To do so, we must first establish some institutional payoffs[30].

We can assign a task to each group, in which the interactions between the agents' strategies and the stochastic environment yield some sort of payoff for the social institution. We consider the institutional payoff to be the sum of the payoffs of all the groups[31]

Each group in each level interacts with the environment and receives a payment for their organization. If a group does not give the right response to the environment[32], the group receives a payoff of $-\eta$ with $\eta > 0$. We analyze the case of $n = 2$, i.e., groups with two agents. As such, we have three kinds of groups[33]:

kind of group	AA	AB	BB
number of groups	$\frac{1}{2}a^2$	ab	$\frac{1}{2}b^2$
group's payoffs	σ	1	σ
failure cost	$-\eta$	$-\eta$	$-\eta$

With $\eta > 0$, $\sigma > -\eta$

We label the AB group the **Mixed Group** and the AA and BB groups the **Homogeneous Groups**. If the group matches the environment[34], we assume that the payoffs

[30] But before proceeding to do so, however, we must consider the agents who are not promoted and will no longer be considered for promotion. We could say that they still exist within the system but are no longer "in the race". In equilibrium, the agents who compete for promotion are included in the same profile at each level. After the selection system has been put into operation, some of the agents are promoted while others are not. If the system is in equilibrium, however, the promoted agent's profile is still identical to that of the agent who has not been promoted. In equilibrium, therefore, in the long run, the entire hierarchy has the same profile, i.e., not only the agents who are "in the race", but also the agents who lose that "race" at any level. This is why, if the promoted agent's profile is efficient, the profile of the agents who are not promoted are so also, provided that the system is in equilibrium.

[31] We consider the group payoff as the social institution payoff. One possible interpretation of the group's payoff is to consider it as the aggregation of their individual performances or payoffs.

[32] None of the agents in that group give the right response to their environment.

[33] $AA(BB)$ groups, composed of two agents who follow the same Strategy $A(B)$, and AB groups, formed by one agent who follows Strategy A and another following Strategy B.

[34] At least one of the agents gives the right answer to the environment in that group.

of the two homogeneous groups are identical, whereas the mixed group's payoff may be different from that of the homogeneous groups. We normalize the mixed group's payoff to one and the homogeneous groups' payoff by the parameter σ (with $\sigma > -\eta$).

If $\sigma > 1$ then the homogeneous groups' payoff is greater than that of the mixed group when both types of groups match the environment. In such a case, if the group that matches the environment is the homogeneous one, (with both agents following the same strategy), there are positive externalities between the two agents. We can imagine the sort of group assignments that are better carried out if the people who must work together have received similar training or follow similar strategies.

If $\sigma < 1$ there are negative externalities between agents following the same strategy when they match the environment. In this case, we can imagine situations in which the task can better done if there is less competition within the group.

If $\sigma = 1$ The heterogeneity in the group does not affect the payoff if the environment is matched.

Therefore, if $\sigma > 1$, heterogeneity in the group has a premium, and if $\sigma < 1$, then homogeneity in the group has a premium provided that the group matches the environment.

The institutional payoff for the population of level t (we have a_t agents following strategy A and b_t agents following strategy B) is:

$$\frac{1}{2}a_t^2(\sigma P1 - (1-P_1)\eta) + \frac{1}{2}b_t^2(\sigma P2 - (1-P_2)\eta) + a_t b_t((P_1+P_2) - (1-P_1-P_2)\eta) \quad (11)$$

The previous expression is quite easy to obtain. For example, the first term is the payoff of the AA groups is obtained as follows: The number of AA groups is $\frac{1}{2}a_t^2$, a proportion P_1 (the probability environment e_1) of them matches the environment and they obtain a payoff of σ each. The remainder do not match the environment, so they get a payoff of $-\eta$. The rest of the terms are obtained analogously.

Since all of the levels have the same environment probabilities, the optimum population profile for one level is also optimum for the other levels as well. We therefore study the optimum population profile for one level [35]. Since $b_t = 1 - a_t$, we have the following institutional payoff function of a level.

$$\Pi(a_t) = \frac{1}{2}(P_2(\eta+\sigma)-\eta) + a_t(P_1(1+\eta) + P_2(1-\sigma)) - a_t^2\frac{1}{2}(P_1+P_2)(2+\eta-\sigma) \quad (12)$$

The following result shows us the optimal population profile.

Proposition 7 *Let us suppose $n = 2$ and the institutional payoff function*(12). *The optimal proportion of agents following strategy A is*[36]*:*

If $P_1 \geq P_2$ then that is:

$$\bar{a} = \begin{cases} \dfrac{P_1(1+\eta)+P_2(1-\sigma)}{(P_1+P_2)(2+\eta-\sigma)} & \textit{for } \sigma < 1 + \frac{P_2}{P_1}(1+\eta) \\ 1 & \textit{for } \sigma \geq 1 + \frac{P_2}{P_1}(1+\eta) \end{cases}$$

[35] Note that the population profile that maximizes the institutional payoff function of a level also maximizes the institutional payoff function of the K levels ($K > 1$).

[36] Supposing $P1 > P2$, note that $(2+\eta-\sigma) = 0$ and $\sigma < 1 + \frac{P2}{P1}(1+\eta)$ are incompatible, because $\sigma = 2+\eta > 1 + \frac{P2}{P1}(1+\eta)$. Therefore, if $\sigma = 2+\eta$ then $\bar{a} = 1$.

If $P_1 < P_2$ then that is:

$$\bar{a} = \begin{cases} \dfrac{P_1(1+\eta)+P_2(1-\sigma)}{(P_1+P_2)(2+\eta-\sigma)} & \text{for } \sigma < 1+\frac{P_1}{P_2}(1+\eta) \\ 0 & \text{for } \sigma \geq 1+\frac{P_1}{P_2}(1+\eta) \end{cases}$$

Note that if the groups that match the environment have the same payoff irrespective of their composition ($\sigma = 1$) then $\bar{a} = \frac{P_1}{P_1+P_2}$ for any η, the optimal population profile does not depend on the cost of failure. If $\sigma = 1$ then the population profile that maximizes the institutional payoff is the same as the one that maximizes the number of groups that match the environment (that occur with $a_t = \frac{P_1}{P_1+P_2}$).

Without loss of generality, now we consider the environment e_1 to be more common, $P_1 > P_2$. As such, the agents who follow A (A-agents) are following the strategy with the higher success rate. The previous proposition shows that there is a threshold of the homogeneous group's payoff $\bar{\sigma} = 1 + \frac{P_2}{P_1}(1+\eta) > 1$. If $\sigma < \bar{\sigma}$, it is optimal to have agents following both strategies, and if $\sigma > \bar{\sigma}$, it is optimal that the entire population only follows the strategy with the higher success rate.

From these results, we obtain the following consequences. Note that $\sigma \in (-\eta, +\infty)$.

Corollary 3 *Let $P_1 > P_2$ and η constant:*
If σ increases then $\bar{a}$ increases, provided that $\sigma < \bar{\sigma} = 1 + \frac{P_2}{P_1}(1+\eta)$.
If σ decreases then $\bar{a}$ decreases, with a minimum value $\bar{a} = \frac{1}{2}$.

An increase in the homogeneous group's payoff (in relation to the mixed group's payoff) makes the number of A-agents (more successful agents) increase in the optimum profile[37] $\bar{a}$. If σ decreases the weight of the A-agents in the optimal profile decreases. However, there is a lower limit beyond which that the optimal profile can never fall [38]. That is $\bar{a} = \frac{1}{2}$.

Corollary 4 *Let $P_1 > P_2$.*
Provided $\sigma < 1$,, if η increases(decreases), then $\bar{a}$ increases(decreases).
Provided $\sigma > 1$, if η increases(decreases), then $\bar{a}$ decreases(increases).
Provided $\sigma = 1$, if η changes, then $\bar{a}$ does not change.

If the homogeneous group's payoff is greater than the mixed group's payoff ($\sigma < 1$) then an increase(decrease) in the failure cost[39] increases (decreases) the weight of the more successful agents in the optimum (with a minimum value[40] to $\bar{a}$, that is $\frac{1}{2}$).

Before understanding that result note that, if $\sigma < 1$ then the optimum profile $\bar{a}$ belongs to the interval[41] $\left(\frac{1}{2}, \frac{P_1}{P_1+P_2}\right)$, see Figure 9.

Note that, if the proportion of agents who follow Strategy A is $a_t = \frac{P_1}{P_1+P_2}$ there is a maximum number of groups matching the environment. However, if $\sigma < 1$, a mixed group (AB) that matches the environment obtains a higher payoff than a homogeneous

[37] Note that $\frac{\partial \bar{a}}{\partial \sigma} = \frac{(P_1-P_2)(1+\eta)}{(P_1+P_2)(2+\eta-\sigma)^2} > 0$

[38] If $P1 > P2$, then we have $\frac{\partial \bar{a}}{\partial \sigma} > 0$ and $\lim_{\sigma \to -\eta} \frac{P1(1+\eta)+P2(1-\sigma)}{(P1+P2)(2+\eta-\sigma)} = \frac{1}{2}$. Note that $\sigma \in (-\eta, +\infty)$.

[39] Note that $\frac{\partial \bar{a}(\eta)}{\partial \eta} = \frac{(P_1-P_2)(1-\sigma)}{(P_1+P_2)(2+\eta-\sigma)^2}$

[40] Note that $\eta \in [0, \infty)$ and if $\sigma < 1$ then $\frac{\partial \bar{a}(\eta)}{\partial \eta} > 0$, therefore $\lim_{\eta \to 0} \bar{a}(\eta) = \lim_{\eta \to 0} \frac{P_1(1+\eta)+P_2(1-\sigma)}{(P_1+P_2)(2+\eta-\sigma)} = \frac{P_1+P_2(1-\sigma)}{(P_1+P_2)(2-\sigma)}$. The previous expresion is increasing in σ. The minimum value that expresion can takes it is for the minimum σ and that is $\sigma = -\eta = 0$, therefore $\left.\frac{P_1+P_2(1-\sigma)}{(P_1+P_2)(2-\sigma)}\right|_{\sigma=0} = \frac{1}{2}$.

[41] If $\sigma < 1$ then $\bar{a} = \frac{P_1(1+\eta)+P_2(1-\sigma)}{(P_1+P_2)(2+\eta-\sigma)}$ and $\lim_{\eta \to +\infty} \frac{P_1(1+\eta)+P_2(1-\sigma)}{(P_1+P_2)(2+\eta-\sigma)} = \frac{P_1}{P_1+P_2}$

group (AA or BB) does so. Therefore, it is possible to increase the total payoff by increasing the number of mixed groups that match the environment. This occurs if a_t decreases, and the total number of groups that match the environment will obviously decrease. In such a case, if the failure cost η increases, then the cost of having too many groups that do not match the environment increases. An increase in η therefore, causes the optimum profile $\bar{a}$ converges to $\frac{P_1}{P_1+P_2}$ in order to decrease the number of groups that do not match the environment.

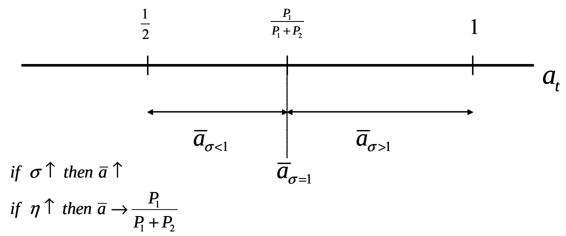


Figure 9: If $\sigma < 1$ then the optimum profile $\bar{a}_{\sigma<1}$ belongs to the interval $\left(\frac{1}{2}, \frac{P_1}{P_1+P_2}\right)$, if $\sigma > 1$ then the optimum profile $\bar{a}_{\sigma>1}$ belongs to the interval $\left(\frac{P_1}{P_1+P_2}, \frac{1}{2}\right)$and if $\sigma = 1$ then the optimum profile $\bar{a}_{\sigma=1}$ is equal to $\frac{P_1}{P_1+P_2}$. Moreover, if σ increases, the optimum profile $\bar{a}$ increases and if the cost of failure increases then the optimum profile converge to $\frac{P_1}{P_1+P_2}$.

If homogeneous and mixed groups' payoffs are equal ($\sigma = 1$) then the failure cost η does not change the optimum profile $\bar{a}$. The system maximizes payoffs by maximizing the number of groups that match the environment, which occurs in $a_t = \frac{P_1}{P_1+P_2}$.

If the mixed group's payoff is smaller than the homogeneous group's payoff ($\sigma > 1$) an increase(decrease) in the failure cost decreases(increases) the weight of the more successful agents in the optimum, with a lower bound[42] for $\bar{a}$, which is $\frac{P_1}{P_1+P_2}$. Note that if $\sigma > 1$ then the optimum profile $\bar{a}$ belongs to the interval[43] $\left(\frac{P_1}{P_1+P_2}, 1\right)$. In this case, a mixed group (AB) that matches the environment obtains a smaller payoff than a homogeneous group (AA or BB) that also matches the environment. We know that in $a_t = \frac{P_1}{P_1+P_2}$ the number of groups matching the environment is maximized. However, if a_t increases then the number of AA groups matching the environment increases and the BB and AB groups matching the environment decrease, but the eventual effect is an increase in the total payoff. The effect of an increase in the failure cost is similar to the case $\sigma < 1$, thus, if η increases the optimum profile approaches to $\bar{a} = \frac{P_1}{P_1+P_2}$, therefore $\bar{a}$ decreases, see 9.

Corollary 5 *Let $P_1 > P_2$. If the "expected" performance gap between the two different kinds of agents increases, i.e., (($P_1 - P_2$) increases), then the presence of the more successful agent in the optimum profile increases, provided that $\sigma < 1 + \frac{P_2}{P_1}(1+\eta)$.*

In conclusion, given P_1 and P_2, (i.e., given the relative advantage of following a given rule), we can obtain any profile between $(\frac{1}{2}, 1)$, as the optimal profile, for the possible values of the parameters σ and η. On the other hand, if we consider the selection

[42] Note that $\eta \in [0, +\infty)$ and if $\sigma > 1$ then $\frac{\partial \bar{a}(\eta)}{\partial \eta} < 0$, therefore $\lim_{\eta\to+\infty} \bar{a}(\eta) = \lim_{\eta\to+\infty} \frac{P1(1+\eta)+P2(1-\sigma)}{(P1+P2)(2+\eta-\sigma)} = \frac{P1}{P1+P2}$

[43] If $\sigma > 1$ then either $\bar{a} = \frac{P_1(1+\eta)+P_2(1-\sigma)}{(P_1+P_2)(2+\eta-\sigma)}$ or $\bar{a} = 1$, and we have that $\frac{\partial \bar{a}(\eta)}{\partial \eta} < 0$ and $\lim_{\eta\to+\infty} \frac{P_1(1+\eta)+P_2(1-\sigma)}{(P_1+P_2)(2+\eta-\sigma)} = \frac{P_1}{P_1+P_2}$

process $S_{[n=2,k,\theta]}$, then changing θ and k (so the stringency degree) we can obtain any global stable equilibrium between the thresholds $(\frac{2P_1-P_2}{P_1+P_2}, 1)$. Note that if $P_1 > P_2$ then $\frac{1}{2} < \frac{2P_1-P_2}{P_1+P_2}$. Thus, not all possible optimum profiles are achieved in this case.

In general, if the optimum is that all, or nearly all, of the population follows the rule with the higher success rate, then the stringency degree of the system must not be too high. However, if the optimum is a heterogeneous profile then the stringency degree must be high.

Obviously, if you consider groups with more than two members, the optimum profile must be in the interval $(\frac{1}{2}, 1)$ (if $P_1 > P_2$). However, as we have shown in section 4 if the stringency degree goes to 1, i.e., the proportion of people promoted is close to 0, then the profile in the long run goes to $\frac{P_1}{P_1+P_2}$. Thus, the interval of possible population profiles in the long run is $\left(\frac{P_1}{P_1+P_2}, 1\right)$. Therefore, as $\frac{1}{2} < \frac{P_1}{P_1+P_2}$, not all possible optimum profiles would be achieved in this case.

Regarding the interesting work Harrington (1998a) which was the main inspiration of the present paper, we could say that his results would change in the case that the degree of stringency increases or decreases. If the level of stringency is increased enough, the stable equilibrium where the entire population following the more successful rigid rule would lose his stability and there will be a heterogeneous population in the long run. Although, we think that the predominance of the rigid rule will hold. And if the degree of stringency is decreasing enough, all of the homogenous equilibria become locally stable equilibria.

On the other hand, we must be careful about the stringency degree of promotion mechanism in hierarchical social systems. As we have shown if we wish to increase the presence in the social system of certain agents with a greater expected success rate, perhaps, in certain contexts, we must decrease the stringency of the promotion instead of increase it.

Appendix

PROOF OF PROPOSITION 1

That proposition is a particular case of the proposition 2.

PROOF OF PROPOSITION 2

We assume $P_1 > P_2$

The dynamics of the selection process $S_{[n=2,k=1,\theta]}$ is given by the equation:

$$a_{t+1} = \begin{cases} \frac{2P_1 a_t - P_1 a_t^2}{P_2 + 2P_1 a_t - (P_1+P_2)a_t^2} & if \quad 2P_1 a_t - (P_1+P_2)a_t^2 + P_2 - \theta \geq 0 \\ \frac{a_t}{\theta}\left(2P_1 - a_t P_1 + \frac{(1-P_1+(-1+a_t)P_2)(-2a_t P_1 - P_2 + a_t^2(P_1+P_2)+\theta)}{1-2a_t P_1 - P_2 + a_t^2(P_1+P_2)}\right) & if \quad 2P_1 a_t - (P_1+P_2)a_t^2 + P_2 - \theta < 0 \end{cases}$$

We can write the previous equation as:

$$a_{t+1} = f(a_t) = \begin{cases} f_1(a_t) & if \quad h(a_t) \geq 0 \\ f_2(a_t) & if \quad h(a_t) < 0 \end{cases}$$

We prove the proposition in several steps.

We name $g(a_t) = f(a_t) - a_t = \begin{cases} g_1(a_t) = f_1(a_t) - a_t & if \quad h(a_t) \geq 0 \\ g_2(a_t) = f_2(a_t) - a_t & if \quad h(a_t) < 0 \end{cases}$

Lemma 1 *The function $g(a_t) = f(a_t) - a_t$ is continuous in $[0, 1]$.*

It is straightforward to show that $f_1(a_t)$ and $f_2(a_t)$ are continuous, both functions are a quotient of polynomials and both denominators are strictly greater than zero. Furthermore, if $2P_1a_t - (P_1 + P_2)a_t^2 + P_2 = \theta$ then $f_1(a_t) = f_2(a_t)$. Since $f(a_t)$ is continuous, $g(a_t) = f(a_t) - a_t$ is continuous as well.■

Lemma 2 *The equation $g(a_t) = 0$ has no more than one root in $(0,1)$.*

It is straightforward to show that the equation $g_1(a_t) = 0$ has the following roots.

$$\bar{a}_1 = 0 \quad \bar{a}_2 = 1 \quad \bar{a}_3 = \tfrac{2P_1-P_2}{P_1+P_2}$$

Now we study the roots of the equation $g_2(a_t) = 0$, first we rewrite the function $g_2(a_t)$.

$$g_2(a_t) = \frac{a_t}{\theta}(2P_1 - a_tP_1 + \tfrac{(1-P_1+(-1+a_t)P_2)(-2a_tP_1-P_2+a_t^2(P_1+P_2)+\theta)}{1-2a_tP_1-P_2+a_t^2(P_1+P_2)}) - a_t$$

$$g_2(a_t) =$$
$$a_t(a_t-1)\tfrac{1}{(1-2a_tP_1-P_2+a_t^2(P_1+P_2))\theta}\big(a_t^2\big(P_2^2-P_1^2\big)+a_t\big((P_1^2-P_2^2)+(P_1-P_2)^2+(P_1+P_2)(1-\theta)\big)+P_2-P_2^2+P_1(P_2+\theta-2)\big)$$

It is straightforward to show that the denominator $\big(1 - 2a_tP_1 - P_2 + a_t^2(P_1 + P_2)\big)\,\theta > 0$.

The equation $g_2(a_t) = 0$ has four roots, two of them are $a_1 = 0$ and $a_2 = 1$, the other two are given by the polynomial:

$$pol[a_t] = a_t^2\big(P_2^2-P_1^2\big)+a_t\big((P_1^2-P_2^2)+(P_1-P_2)^2+(P_1+P_2)(1-\theta)\big)+\big(P_2-P_2^2+P_1(P_2+\theta-2)\big)$$

The previous polynomial determine the roots a_3, a_4 and also the sign of the function $g_2(a_t)$. If $pol[a_t] < 0$ then $g_2(a_t) > 0$ (thus, the population increase at that point, provided that $h(a_t) < 0$).

Let $pol[a_t] = \alpha a_t^2+\beta a_t+\gamma$ with (it is straightforward to show the sign of the following expressions):

$$\begin{aligned}
&\alpha = \big(P_2^2 - P_1^2\big) < 0 \\
&\beta = \Big((P_1^2 - P_2^2) + (P_1 - P_2)^2 + (P_1 + P_2)\,(1 - \theta)\Big) > 0 \\
&\gamma = P_2 - P_2^2 + P_1\,(P_2 + \theta - 2) < 0
\end{aligned} \tag{13}$$

Provided that $\beta^2 - 4\alpha\gamma > 0$ we will have two different real roots from the $pol[a_t]$. Furthermore, since $\alpha < 0$ and $\beta > 0$ we have that:

$$a_3 = \tfrac{-\beta+\sqrt{\beta^2-4\alpha\gamma}}{2\alpha} < \tfrac{-\beta-\sqrt{\beta^2-4\alpha\gamma}}{2\alpha} = a_4$$

Now we show that if a_4 is real root then it is greater or equal than 1. $(a_4 \geq 1)$.

$a_4 = \frac{-\beta-\sqrt{\beta^2-4\alpha\gamma}}{2\alpha} < 1 \Leftrightarrow -\beta - \sqrt{\beta^2 - 4\alpha\gamma} > 2\alpha \Leftrightarrow$

$\Leftrightarrow -\sqrt{\beta^2 - 4\alpha\gamma} > 2\alpha + \beta$

Note that, if $a_4 < 1$ then it is necessary that $2\alpha + \beta < 0$. Next we show that it is no possible to hold simultaneously both conditions; $\beta^2 - 4\alpha\gamma > 0$ (real a_4) and $2\alpha + \beta < 0$.

Using (13) it is straightforward to show that:

$2\alpha + \beta < 0 \Leftrightarrow (P_1 + P_2)\,(1 - \theta) - 2P_2\,(P_1 - P_2) < 0$

$$\beta^2 - 4\alpha\gamma > 0 \Leftrightarrow$$
$$-4\left(P_1 - P_2\right)^2 \left(\left(P_1 + P_2\right) + P_1P_2 - \left(P_1 + P_2\right)^2\right) + \left(P_1 + P_2\right)^2 \left(-1 + \theta\right)^2 > 0$$

It follows that:

$2\alpha + \beta < 0 \Leftrightarrow \left(P_1 + P_2\right)\left(1 - \theta\right) < 2P_2\left(P_1 - P_2\right)$(since the two terms are positive) $\Leftrightarrow \left(P_1 + P_2\right)^2\left(1 - \theta\right)^2 < \left(2P_2\left(P_1 - P_2\right)\right)^2$.

$$\beta^2 - 4\alpha\gamma > 0 \Leftrightarrow \left(P_1 + P_2\right)^2\left(-1 + \theta\right)^2 > 4\left(P_1 - P_2\right)^2\left(\left(P_1 + P_2\right) + P_1P_2 - \left(P_1 + P_2\right)^2\right)$$

Therefore it necessary that:

$4\left(P_1 - P_2\right)^2\left(\left(P_1 + P_2\right) + P_1P_2 - \left(P_1 + P_2\right)^2\right) < \left(2P_2\left(P_1 - P_2\right)\right)^2 \Leftrightarrow$

$P_1^2 - P_1\left(1 - P_2\right) + P_2\left(2P_2 - 1\right) > 0 \Leftrightarrow$

$P_1(-1 + P_1) + P_2(-1 + P_1) + P_2^2 + P_2^2 > 0$

It is straightforward to show that the previous expression is always smaller than zero, for that it is no possible simultaneously conditions; $\beta^2 - 4\alpha\gamma > 0$ and $2\alpha + \beta < 0$, thus a_4 can not be smaller than one. Therefore we can say that $a_4 \geq 1$.

We know that, if the equation $g_1(a_t) = 0$ has one root in $(0, 1)$, then it has to be $\bar{a}_3$ and if $g_2(a_t) = 0$ has one root in $(0, 1)$ then that has to be a_3. Therefore the equation $g(a_t) = 0$ can have zero, one or two roots in $(0, 1)$. We have only two candidates to be roots in $(0, 1)$ of the equation $g(a_t) = 0$ they are $\bar{a}_3$ and a_3.

Now we prove that the equation $g(a_t) = 0$ can not have two roots in $(0, 1)$.

The explicit expressions to $\bar{a}_3$ and a_3 are:

$\bar{a}_3 = \frac{2P_1 - P_2}{P_1 + P_2}$

$$a_3 = \frac{-\left(\left(P_1^2 - P_2^2\right) + \left(P_1 - P_2\right)^2 + \left(P_1 + P_2\right)(1-\theta)\right) + \sqrt{\left(\left(P_1^2 - P_2^2\right) + \left(P_1 - P_2\right)^2 + \left(P_1 + P_2\right)(1-\theta)\right)^2 - 4\left(P_2^2 - P_1^2\right)\left(P_2 - P_2^2 + P_1\left(P_2 + \theta - 2\right)\right)}}{2\left(P_2^2 - P_1^2\right)}$$

Remember that $g(a_t) = \begin{cases} g_1(a_t) & if \quad h(a_t) \geq 0 \\ g_2(a_t) & if \quad h(a_t) < 0 \end{cases}$.

First, we will show that $h(\bar{a}_3) \geq 0 \Leftrightarrow \theta \leq \frac{3P_1P_2}{P_1 + P_2}$.

$$h(\bar{a}_3) = 2P_1\bar{a}_3 - (P_1 + P_2)\bar{a}_3^2 + P_2 - \theta = 2P_1\frac{2P_1 - P_2}{P_1 + P_2} - (P_1 + P_2)\left(\frac{2P_1 - P_2}{P_1 + P_2}\right)^2 + P_2 - \theta \geq 0 \Leftrightarrow$$
$$\Leftrightarrow \frac{3P_1P_2}{P_1 + P_2} - \theta \geq 0.$$

After that, we will show that $h(a_3) < 0 \Rightarrow \theta > \frac{3P_1P_2}{P_1 + P_2}$.

$h(a_3) < 0 \Rightarrow$

$$2P_1^2 + P_1(\theta - 1 - 4P_2) + P_2(\theta - 1 + 2P_2) + \sqrt{\left(\left(P_1^2 - P_2^2\right) + \left(P_1 - P_2\right)^2 + \left(P_1 + P_2\right)(1-\theta)\right)^2 - 4\left(P_2^2 - P_1^2\right)\left(P_2 - P_2^2 + P_1\left(P_2 + \theta - 2\right)\right)} < 0$$
$$\Rightarrow *$$

(Note that $2P_1^2 + P_1\left(\theta - 1 - 4P_2\right) + P_2\left(\theta - 1 + 2P_2\right)$ smaller than zero is a necessary condition to get $h(a_3) < 0$. Therefore, $-\left(2P_1^2 + P_1(\theta - 1 - 4P_2) + P_2(\theta - 1 + 2P_2)\right) > 0$)

$$* \Rightarrow$$
$$\left(\left(P_1^2 - P_2^2\right)\left(P_1 - P_2\right)^2 + \left(P_1 + P_2\right)(1-\theta)\right)^2 - 4\left(P_2^2 - P_1^2\right)\left(P_2 - P_2^2 + P_1\left(P_2 - \theta\right)\right) < \left(-\left(2P_1^2 + P_1(\theta - 1 - 4P_2) + P_2(\theta - 1 + 2P_2)\right)\right)^2$$
$$\Rightarrow \frac{3P_1P_2}{P_1 + P_2} - \theta < 0 \blacksquare$$

Remark 1 *We can state that the equation $g(a_t)=0$ has no more than one root in $(0,1)$. If $\theta \leq \frac{3P_1P_2}{P_1+P_2}$ then $\bar{a}_3$ is the only possibility. If $\theta > \frac{3P_1P_2}{P_1+P_2}$ then a_3 is the only possibility.*

Lemma 3 $g'(0) > 0$

It is straightforward to show that $g_1'(0) = \frac{2P_1}{P_2} - 1$, since $P_1 > P_2$ we can state that $g_1'(0) > 0$.

We can obtain that $g_2'(0) = \frac{P_2 - P_2^2 + P_1(-2+P_2+\theta)}{(-1+P_2)\theta}$, the denominator is negative $(-1+P_2)\theta < 0$, and it is easy to show that the numerator is negative as well, $P_2 - P_2^2 + P_1(-2+P_2+\theta) = P_2(1-P_2) + P_1(P_2 + (\theta - 2)) < 0$, therefore $g'(0) > 0$.$\blacksquare$

Lemma 4 *a) If $\theta \leq P_1 \Rightarrow \left[g'(1) \gtreqless 0 \Longleftrightarrow 2P_2 \gtreqless P_1\right]$*
b) If $\theta > P_1 \Rightarrow \left[g'(1) \gtreqless 0 \Longleftrightarrow \theta \lesseqgtr \frac{P_1^2 + 2P_2 - P_1(1+P_2)}{P_2}\right]$

First, note that:
$g'(1) = g_1'(1)$ provided that $h(1) \geq 0$ and $h(1) \geq 0 \Leftrightarrow P_1 - \theta \geq 0$
$g'(1) = g_2'(1)$ provided that $h(1) < 0$ and $h(1) < 0 \Leftrightarrow P_1 - \theta < 0$
The part *a*) is easy to prove:
If $\theta \leq P_1$ then $g'(1) = g_1'(1) = \frac{2P_2}{P1} - 1$, that expression is greater, equal or smaller than zero if $2P_2 \gtreqless P_1$.
The part *b*):
If $\theta > P_1$ then $g'(1) = g_2'(1) = \frac{-P_1^2 + P_1(1+P_2) + P_2(-2+\theta)}{(-1+P_1)\theta}$, the denominator is negative $(-1+P_1)\theta < 0$, the numerator is negative positive or zero $-P_1^2 + P_1(1+P_2) + P_2(-2+\theta) \lesseqgtr 0$ provide $\theta \lesseqgtr \frac{P_1^2 + 2P_2 - P_1(1+P_2)}{P_2}$. Therefore $g'(1) \gtreqless 0$ if $\theta \lesseqgtr \frac{P_1^2+2P_2-P_1(1+P_2)}{P_2}$.$\blacksquare$

The next lemma arranges the thresholds of the parameter θ.

Lemma 5 *If $2P_2 > (\leq) P_1 \Rightarrow \left[P_1 < (\geq) \frac{3P_1P_2}{P_1+P_2} < (\geq) \frac{P_1^2+2P_2-P_1(1+P_2)}{P_2}\right]$*

First, it is straightforward to show that:
$P_1 < (\geq) \frac{3P_1P_2}{P_1+P_2} \Leftrightarrow P_1(P_1 - 2P_2) < (\geq) 0 \Leftrightarrow 2P_2 > (\leq) P_1$
Next,

$$\frac{3P_1P_2}{P_1+P_2} < (\geq) \frac{P_1^2+2P_2-P_1(1+P_2)}{P_2} \Leftrightarrow (P_1 - 2P_2)(P_2 + P_1(1 - P_1 - 2P_2)) < (\geq) 0$$

Since $(P_2 + P_1(1 - P_1 - 2P_2) > 0$ the sign depend on $(P_1 - 2P_2)$.$\blacksquare$

Lemma 6 *a_3 is increasing in θ and $a_3 = 1$ with $\theta = \frac{P_1^2+2P_2-P_1(1+P_2)}{P_2}$*

The expression of root a_3 is:

$$a_{3=}$$

$$\frac{-\left((P_1^2-P_2^2)+(P_1-P_2)^2+(P_1+P_2)(1-\theta)\right)+\sqrt{\left((P_1^2-P_2^2)+(P_1-P_2)^2+(P_1+P_2)(1-\theta)\right)^2+4\left(P_1^2-P_2^2\right)(P_2-P_2^2+P_1(P_2+\theta-2))}}{2\left(P_2^2-P_1^2\right)} \tag{14}$$

We calculate $\dfrac{\partial\, a_3}{\partial\, \theta}$ and that is equal to the following expression:

$$\left(\frac{(P_1+P_2)(1-\theta)}{\sqrt{\left((P_1^2-P_2^2)+(P_1-P_2)^2+(P_1+P_2)(1-\theta)\right)^2+4\left(P_1^2-P_2^2\right)(P_2-P_2^2+P_1(P_2+\theta-2))}} - 1 \right) \frac{1}{2(P_1^2-P_2^2)}$$

Therefore, if:

$$\frac{(P_1+P_2)(1-\theta)}{\sqrt{\left((P_1^2-P_2^2)+(P_1-P_2)^2+(P_1+P_2)(1-\theta)\right)^2+4\left(P_1^2-P_2^2\right)(P_2-P_2^2+P_1(P_2+\theta-2))}} > 1 \tag{15}$$

then we can state that $\dfrac{\partial\, a_3}{\partial\, \theta} > 0$.

Note that $\left((P_1^2-P_2^2)+(P_1-P_2)^2+(P_1+P_2)(1-\theta)\right)^2+4\left(P_1^2-P_2^2\right)(P_2-P_2^2+P_1(P_2+\theta-2))$ must be greater or equal than zero, if it is not, then a_3 is not real, see 14. Moreover, if that expression is equal to zero then $a_3 = a_4 > 1$ (in that case $a_3 = a_4 = \frac{-\beta}{2\alpha} > 1$).

Therefore we can state that (15) is true if:
$\left((P_1^2-P_2^2)+(P_1-P_2)^2+(P_1+P_2)(1-\theta)\right)^2+4\left(P_1^2-P_2^2\right)(P_2-P_2^2+P_1(P_2+\theta-2)) < (P_1+P_2)^2(1-\theta)^2$
$\Leftrightarrow\ 4(P1-P2)^2(P1^2+P1(-1+P2)+(-1+P2)P2) < 0$

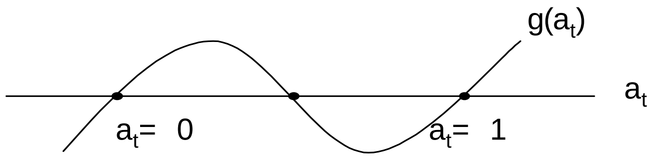
It is straightforward to show that the expression $P1^2 + P1(-1+P2) + (-1+P2)P2$ is smaller than zero, therefore we can state that $\dfrac{\partial\, a_3}{\partial\, \theta} > 0$.

On the other hand, if we do $\theta = \frac{P_1^2+2P_2-P_1(1+P_2)}{P_2}$ in the expression (14) then it is straightforward to show that $a_3 = 1$.■

We have the following results:

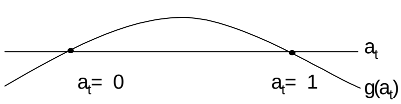
The function $g(a_t)$ is continuous in $[0,1]$ and $g'(0) > 0$, the equation $g(a_t) = 0$ has no more than one root in $(0,1)$. Therefore we have the following three possibilities:

1) If $g'(1) > 0$ then the only possibility is (for the function $g(a_t)$):

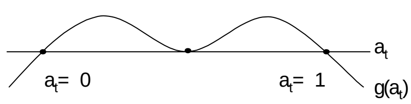


In such a case, the system converges to the only root in $(0,1)$, at which point the system has a global equilibrium. This is true provided that there are not periodic points, we prove that statement at the end of this prove.

2) If $g'(1) < 0$ then we could have only two cases:



Case 1



Case 2

In Case 1, the system converge to $a_t = 1$, this point is a global equilibrium of the system. Case 2 is not possible, we shall now prove this statement.

We know that either $\bar{a}_3$ or a_3 is the only possible root in $(0,1)$. We shall now prove that if $g'(1) \leq 0$ then the equation $g(a) = 0$ does not have any roots in $(0,1)$. Note that by *Lemma* 4 :

- If $\theta \leq P_1$ and $g'(1) \leq 0$ then by *Lemma* 4 necessarily $2P_2 \leq P_1$ and it is straightforward to show that if $2P_2 \leq P_1$ then $\bar{a}_3 \geq 1$,and it can not be the root. We shall now show that if $2P_2 \leq P_1$ then a_3 can not be root of $g(a) = 0$. The following Figure 10 represents the previous lemmas if $2P_2 < P_1$.

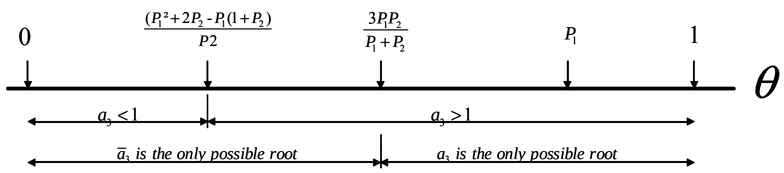


Figure 10: The line represents the parameter θ and we indicate the different thresholds of θ on it. Below the line, we show what happens when the parameter θ is inside those intervals.

Therefore, if $2P_2 < P_1$, then the equation $g(a) = 0$ can not have any root in $(0, 1)$, and Case 2 is not possible, even if $\theta > P_1$.

Note that if $2P_2 = P_1$ then $\frac{P_1^2+2P_2-P_1(1+P_2)}{P_2} = \frac{3P_1P_2}{P_1+P_2} = P_1$ and we have that $\bar{a}_3 = 1$ and a_3 can not be the root of the equation $g(a) = 0$.

- We now consider the case of $\theta > P_1$ and $2P_2 > P_1$. By *Lemma* 4 we know that if $\theta > P_1$ and $g'(1) < 0$ then necessarily $\theta > \frac{P_1^2+2P_2-P_1(1+P_2)}{P_2}$ and it is straightforward to show that if $\theta > \frac{P_1^2+2P_2-P_1(1+P_2)}{P_2}$ then $a_3 > 1$ (if $\theta = \frac{P_1^2+2P_2-P_1(1+P_2)}{P_2}$ then $a_3 = 1$ and a_3 is increasing in θ). Next we show in Figure 11 that if $\theta > \frac{P_1^2+2P_2-P_1(1+P_2)}{P_2}$ and $2P_2 > P_1$ then $\bar{a}_3$ can not be root of $g(a) = 0$.

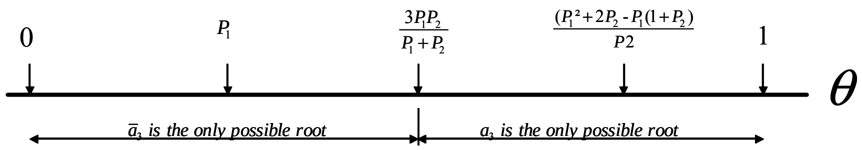
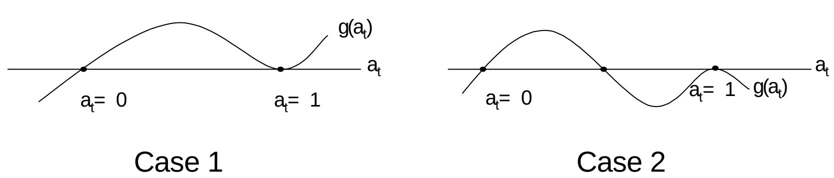


Figure 11: The line represents the parameter θ and we indicate the different thresholds of θ on it. Below the line, we show what happens when the parameter θ is inside those intervals.

Therefore, if $g'(1) < 0$ then the equation $g(a) = 0$ does not have roots in $(0, 1)$.

3) If $g'(1) = 0$ there could be two different cases:



In Case 1, the system converges to $a_t = 1$, which is the point of the global equilibrium of the system. Case 2 is not possible that statement is already proved previously.

In brief:

- If $g'(1) > 0$ then the equation $g(a_t) = 0$ has one root in $(0,1)$, that point is a stable globally equilibrium of the system. We name that point $a^*[S_{[n=2,k=1,\theta]}]$ The only possibilities are either $a^*[S_{[n=2,k=1,\theta]}] = \bar{a}_3$ (if $\theta \leq \frac{3P_1P_2}{P_1+P_2}$) or $a^*[S_{[n=2,k=1,\theta]}] = a_3$ (if $\theta > \frac{3P_1P_2}{P_1+P_2}$).

- If $g'(1) \leq 0$ then the equation $g(a_t) = 0$ has none root in $(0,1)$, therefore $a^*[S_{[n=2,k=1,\theta]}] = 1$.

- By *Lemma* 4 if $2P_2 > P_1$ then we have that $g'(1) > 0$ provided that $\theta \in (0, \frac{P_1^2+2P_2-P_1(1+P_2)}{P_2})$. Therefore, $a^*[S_{[n=2,k=1,\theta]}] = \bar{a}_3$ provided that $\theta \in (0, \frac{3P_1P_2}{P_1+P_2})$, $a^*[S_{[n=2,k=1,\theta]}] = a_3$ provided that $\theta \in (\frac{3P_1P_2}{P_1+P_2}, \frac{P_1^2+2P_2-P_1(1+P_2)}{P_2})$ and $a^*[S_{[n=2,k=1,\theta]}] = 1$ provided that $\theta > \frac{P_1^2+2P_2-P_1(1+P_2)}{P_2}$.

We know that if $2P_2 \leq P_1$ then $g'(1) \leq 0$. In such a case, $a^*[S_{[n=2,k=1,\theta]}] = 1$ and the value of θ is not important at all.

We shall now prove that there are not periodic points[44] *of* $a_{t+1} = f(a_t)$.

The function $f(a_t) = \begin{cases} f_1(a_t) & if \quad h(a_t) \geq 0 \\ f_2(a_t) & if \quad h(a_t) < 0 \end{cases}$ either has just one equilibrium point in $(0,1)$ or none at all. If it has none, there are obviously no periodical points, and in the case of the function having just one equilibrium point in $(0,1)$, if the function is increasing from values that are higher than the inner equilibrium point, any possibility of there being periodical points completely would disappears. We shall now prove that the function $f(a_t)$ is increasing from values that are higher than the inner equilibrium point.

The function $f_1(a_t)$ is increasing in a_t:

$$f_1(a_t) = \frac{2P1a_t - P1a_t^2}{P_2+2P1a_t-(P_1+P_2)a_t^2}; f_1'(a_t) = \frac{(1-a_t+a_t^2)P_1P_2}{\left(P_2+2P1a_t-(P_1+P_2)a_t^2\right)^2} > 0$$

Lemma 7 *The function* $f_2(a_t)$ *is increasing in* $a_t \in [\frac{P_1}{P_1+P_2}, 1]$

$$f_2(a_t) = \frac{a_t}{\theta}(2P_1 - a_tP_1 + (1 - P_1 + (-1+a_t)P_2)\frac{(-2a_tP_1 - P_2 + a_t^2(P_1+P_2)+\theta))}{1-2a_tP_1-P_2+a_t^2(P_1+P_2)})$$

Let $w(a_t) = P_2 + 2P1a_t - (P_1+P_2)a_t^2$, that is a concave function with a maximum in $a_t = \frac{P_1}{P_1+P_2}$. Note that $h(a_t) = w(a_t) - \theta$, therefore $f(a_t) = f_2(a_t)$ provided that $w(a_t) < \theta$. We can rearrange $f_2(a_t)$ in the following way:

$$f_2(a_t) = \frac{1}{\theta}(2P_1a_t - P_1a_t^2 + ((1-P_1-P_2)\,a_t + P_2a_t^2)\frac{\theta - w(a_t)}{1-w(a_t)})$$

$$f_2'(a_t) = \frac{1}{\theta}\,(2P_1 - P_1a_t)((1-P_1-P_2)+P_2a_t)\,\frac{\theta-w(a_t)}{1-w(a_t)}) + ((1-P_1-P_2)a_t+P_2a_t^2)\,\frac{w'(a_t)(\theta-1)}{(1-w(a_t))^2}$$

[44] The point b is called a *periodic point* of $x_{t+1} = f(x_t)$ if for some positive integer k, $f^k(b) = b$, See Elaydi [1996].

It should be noted that all of the terms in the previous expression are always positive, except for $(\theta-1)$, which is always negative and $w'(a_t)$, which may be either positive or negative. The function $w(a_t)$ is concave, with a maximum in $a_t=\frac{P_1}{P_1+P_2}$. Therefore $w'(a_t)<0$ provided that $a_t>\frac{P_1}{P_1+P_2}$■

Note that a_3 is increasing in θ and, if $\theta=\frac{3P_1P_2}{P_1+P_2}$, then $a_3=\bar{a}_3$, we know that if $\theta\leq\frac{3P_1P_2}{P_1+P_2}$ then a_3 can not be an equilibrium point. Therefore, the minimum value of an equilibrium point belong to $(0,1)$ is $\frac{2P_1-P_2}{P_1+P_2}$.

To complete the proof, we need only show that $\frac{P_1}{P_1+P_2}<\frac{2P_1-P_2}{P_1+P_2}$, therefore we can state that the function $f(a_t)$ does not have periodic points.■

Corollaries 1 and 2 have been proved in the previous proof.

PROOF OF PROPOSITION 3

We assume $P_1>P_2$

The dynamics of the selection process $S_{[n=2,k=2,\theta]}$ is given by the equation:

$$a_{t+1}=\begin{cases}\frac{a_tP_1}{a_t(P_1-P_2)+P_2} & if \quad a_t(P_1-P_2)+P_2\geq\theta\\ \frac{a_t(P_1-P_2+(1-P_1)\theta)+a_t^2(P_1-P_2)}{(1-a_t(P_1-P_2)-P_2)\theta} & if \quad a_t(P_1-P_2)+P_2<\theta\end{cases}$$

We can write the previous equation as:

$$a_{t+1}=f(a_t)=\begin{cases}f_1(a_t) & if \quad h(a_t)\geq 0\\ f_2(a_t) & if \quad h(a_t)<0\end{cases}$$

We name $g(a_t)=f(a_t)-a_t=\begin{cases}g_1(a_t)=f_1(a_t)-a_t & if \quad h(a_t)\geq 0\\ g_2(a_t)=f_2(a_t)-a_t & if \quad h(a_t)<0\end{cases}$

We shall now calculate the stationary points:

The roots of the equation $g_1(a_t)=0$ are $\bar{a}_1=0$ and $\bar{a}_2=1$

$$g_1(a_t)=\frac{a_tP_1}{a_t(P_1-P_2)+P_2}-a_t=\frac{a_t\left(P_1-P_2\right)(1-a_t)}{a_t(P_1-P_2)+P_2}$$

Moreover, $g_1(a_t)>0$ provided that $a_t\in(0,1)$

The roots of the equation $g_2(a_t)=0$ are $a_1=0$ and $a_2=1$

$$g_2(a_t)=\frac{a_t(P_1-P_2+(1-P_1)\theta)+a_t^2(P_1-P_2)}{(1-a_t(P_1-P_2)-P_2)\theta}-a_t=\frac{a_t\left(P_1-P_2\right)(1-a_t)(1-\theta)}{(1-a_t(P_1-P_2)-P_2)\theta}$$

Moreover, it is straightforward to show that $g_1(a_t)>0$ provided that $a_t\in(0,1)$.

Therefore, $a_1=0$ and $a_2=1$ are the only stationary points and $g(a_t)>0$ for all $a_t\in(0,1)$, which means that the system converge in the long run to $a=1$.

$$a^*\left[S_{[n=2,k=2,\theta]}\right]=1$$

■

PROOF OF PROPOSITION 4

Without loss of generality we consider $P_1>P_2$.

The dynamics of the selection process $S_{[n,k=1,\theta=1]}$ is given by the equation:

$$a_{t+1}=(1-(1-a_t)^n)P_1+a_t^nP_2+a_t(1-P_1-P_2)$$

Let $a_{t+1}=f(a_t)$ and $g(a_t)=f(a_t)-a_t$

We study the function $g(a_t)$.

The stationary points are the solution of this equation $g(a_t)=0$.

$$(1-(1-a_t)^n)P_1 + a_t^n P_2 - a_t(P_1+P_2) = 0$$

The previous equation does not have an explicit solution. We study the first and the second derivatives.

$$g'(a_t) = n\,(1-a_t)^{n-1}P_1 + n\,a_t^{n-1}P_2 - (P_1+P_2)$$
$$g''(a_t) = -n\,(n-1)\,(1-a_t)^{n-2}P_1 + n\,(n-1)\,a_t^{n-2}P_2$$

We now calculate the inflection points $g''(a_t) = 0$ in $a_t \in (0,1)$.

$$-n\,(n-1)\,(1-\hat{a})^{n-2}P_1 + n\,(n-1)\,\hat{a}P_2 = 0$$
$$(1-\hat{a})^{n-2}P_1 = \hat{a}P_2$$

As $(1-\hat{a}), \hat{a}, P_1, P_2 > 0$ we can write:

$$(1-\hat{a})P_1^{\frac{1}{n-2}} = \hat{a}P_2^{\frac{1}{n-2}}$$

$$\hat{a} = \frac{P_1^{\frac{1}{n-2}}}{P_1^{\frac{1}{n-2}} + P_2^{\frac{1}{n-2}}} = \frac{1}{1+(\frac{P_2}{P_1})^{\frac{1}{n-2}}}$$

Therefore, if $n > 3$ we can then state that the function $g(a_t)$ has no more than one inflection point in $a_t \in (0,1)$. (Note that $\hat{a}$ must be smaller than one, thus, $(\frac{P_2}{P_1})^{\frac{1}{n-2}}$ must be positive).

Obviously, $a = 0$ and $a = 1$ are roots of the equation $g(a_t) = 0$. We calculate the second derivative in these points.

$$g''(0) = -n(n-1)P_1 < 0 \qquad\qquad g''(1) = n(n-1)P_2 > 0$$

Therefore, the function $g(a_t)$ is concave in $a_t = 0$ and convex in $a_t = 1$. As $g(a_t)$ is continuous in $a_t \in [0,1]$, the function $g(a_t)$ is concave in $a_t \in [0,\hat{a})$ and convex in $a_t \in (\hat{a},1]$.

We study the first derivative of $g(a_t)$ in $a = 0$ and $a = 1$:

$$g'(0) = (n-1)P_1 - P_2 > 0$$

Thus, the function $g(a_t)$ in $a = 0$ is increasing.

$$g'(1) = (n-1)P_2 - P_1$$

Thus, the function $g(a_t)$ in $a = 1$ is:

$$\text{increasing if } (n-1)P_2 > P_1$$
$$\text{decreasing if } (n-1)P_2 < P_1$$

Therefore, if $(n-1)P_2 > P_1$, the function $g(a_t)$ is necessarily equal to zero at just one point in the open interval $(0,1)$. Calling that point a^*_{M1}, we can state that the function $g(a_t)$ is greater than zero in $a_t \in (0, a^*_{M1})$, and smaller that zero in $a_t \in (a^*_{M1}, 1)$. Thus, $a^*\left[S_{[n,k=1,\theta=1]}\right] = a^*_{M1}$ if $(n-1)P_2 > P_1$.

On the other hand, if $(n-1)P_2 \le P_1$ then the function $g(a_t)$ can not have a value of zero in the interval $(0,1)$. Moreover, $g(a_t)$ will be positive in $a_t \in (0,1)$, thus $a^*\left[S_{[n,k=1,\theta=1]}\right] = 1$ if $(n-1)P_2 \le P_1$.

To complete the proof we need only note that periodical points are not possible. This is because the first derivative of the function $f(a_t)$ is positive and, as such, $f(a_t)$ is increasing. Periodical points are therefore not possible.

$$f(a_t) = (1-(1-a_t)^n)P_1 + a_t^n P_2 + a_t(1-P_1-P_2)$$
$$f'(a_t) = n\,(1-a_t)^{n-1}P_1 + n\,a_t^{n-1}P_2 + (1-P_1-P_2) > 0 \quad \blacksquare$$

PROOF OF COROLLARY 3

We begin by maintaining all of the previous assumptions for the proof of Proposition 4.

We assume that $(n-1)P_2 > P_1$, it implies that $a^*_{M1} \in (0,1)$

Let $g(a;\,\bar{n})$ be the function $g(a_t)$ evaluates in $a_t = a$, with a parameter $n = \bar{n}$.

Let $g(\tilde{a};\,\bar{n}) = 0$

Let $g(\breve{a};\,\bar{n}+1) = 0$

We wish to prove that $\breve{a} < \tilde{a}$ (provided that $(\bar{n}-1)P_2 > P_1$)

Note that proving the previous statement is the same as proving that $g(\tilde{a};\bar{n}+1) < 0$ (Remember the characteristics of the function $g(a_t)$)

Therefore, we have to prove that $g(\tilde{a};\bar{n}+1) < 0$.

We know that:

$$g(\tilde{a};\,\bar{n}) = (1-(1-\tilde{a})^{\bar{n}})P_1 + \tilde{a}^{\bar{n}}P_2 - \tilde{a}(P_1+P_2) = 0 \Leftrightarrow$$

$$(1-\tilde{a})P_1 - (1-\tilde{a})^{\bar{n}}P_1 + \tilde{a}^{\bar{n}}P_2 - \tilde{a}P_2 = 0 \Leftrightarrow \tag{16}$$
$$\frac{\tilde{a}-\tilde{a}^{\bar{n}}}{(1-\tilde{a})-(1-\tilde{a})^{\bar{n}}} = \frac{P_1}{P_2}$$

On the other hand, if:

$$g(\tilde{a};\,\bar{n}+1) = (1-(1-\tilde{a})^{\bar{n}+1})P_1 + \tilde{a}^{\bar{n}+1}P_2 - \tilde{a}(P_1+P_2) < 0 \Leftrightarrow$$
$$(1-\tilde{a})P_1 - (1-\tilde{a})^{\bar{n}+1}P_1 + \tilde{a}^{\bar{n}+1}P_2 - \tilde{a}P_2 < 0$$

Given(16) , the previous expression will be true if and only if:

$$\tilde{a}^{\bar{n}}P_2 - \tilde{a}^{\bar{n}+1}P_2 > (1-\tilde{a})^{\bar{n}}P_1 - (1-\tilde{a})^{\bar{n}+1P_1}P_1 \Leftrightarrow$$
$$\frac{\tilde{a}^{\bar{n}} - \tilde{a}^{\bar{n}+1}}{(1-\tilde{a})^{\bar{n}} - (1-\tilde{a})^{\bar{n}+1}} > \frac{P_1}{P_2}\left(= \frac{\tilde{a}-\tilde{a}^{\bar{n}}}{(1-\tilde{a})-(1-\tilde{a})^{\bar{n}}}\right) \Leftrightarrow$$
$$\frac{\tilde{a}^{\bar{n}} - \tilde{a}^{\bar{n}+1}}{(1-\tilde{a})^{\bar{n}} - (1-\tilde{a})^{\bar{n}+1}} > \frac{\tilde{a}-\tilde{a}^{\bar{n}}}{(1-\tilde{a})-(1-\tilde{a})^{\bar{n}}} \Leftrightarrow$$
$$\frac{\tilde{a}^{\bar{n}-2}}{(1-\tilde{a})^{\bar{n}-2}} > \frac{1-\tilde{a}^{\bar{n}-1}}{1-(1-\tilde{a})^{\bar{n}-1}} \Leftrightarrow$$
$$\frac{\tilde{a}^{\bar{n}-2}}{1-\tilde{a}^{\bar{n}-1}} > \frac{(1-\tilde{a})^{\bar{n}-2}}{1-(1-\tilde{a})^{\bar{n}-1}}$$

It is straightforward to show that the left hand of the previous expression is increasing in $\tilde{a}$, and the right hand is decreasing in $\tilde{a}$. Moreover, if $\tilde{a} = \frac{1}{2}$ the two term are equal. Therefore, if $\tilde{a} > \frac{1}{2}$ then the previous expression will be true.

Note that:

$$g(\frac{1}{2};\,\bar{n}) = (1-(1-\frac{1}{2})^{\bar{n}})P_1 + \left(\frac{1}{2}\right)^{\bar{n}}P_2 - \frac{1}{2}(P_1+P_2) =$$
$$\left(\frac{1}{2} - \frac{1}{2^{\bar{n}}}\right)(P_1 - P_2)$$

Since the previous expression is greater than zero, we can state that $\tilde{a} > \frac{1}{2}$ ■

PROOF OF PROPOSITION 5

The dynamics of the selection process $S_{[n,k=1,\theta\leq\hat{\theta}]}$ is given by the equation:

$$a_{t+1} = \frac{(1-(1-a_t)^n)P_1}{(1-(1-a_t)^n)P_1 + (1-a_t^n)P_2}$$

Let be $a_{t+1} = f(a_t)$ and

$$g(a_t) = f(a_t) - a_t = \frac{(1-(1-a_t)^n)P_1}{(1-(1-a_t)^n)P_1+(1-a_t^n)P_2} - a_t =$$
$$\frac{P_1-P_1(1-a_t)^n-a_tP_1+a_tP_1(1-a_t)^n-a_tP_2+a_t^{1+n}P_2}{P_1+P_2-P_1(1-a_t)^n-P_2a_t^n} =$$
$$\frac{(1-(1-a_t)^{n+1})P_1+a_t^{n+1}P_2-a_t(P_1+P_2)}{P_1+P_2-P_1(1-a_t)^n-P_2a_t^n}$$

Since the denominator is greater than zero, the sign and the roots of the previous expression depend on the numerator. Notice that the only difference between the function $g(a_t)$ in the previous proposition 4 and this numerator is the parameter n. We can therefore state that:

$$a^*[S_{[n+1,1,\theta=1]}] = a^*[S_{[n,1,\theta\leq\hat{\theta}]}]$$

■

Moreover, since the equilibrium is decreasing in n (provided that $P_1 > P_2$):

$$a^*[S_{[n,1,\theta=1]}] \geq a^*[S_{[n,1,\theta\leq\hat{\theta}]}]$$

PROOF OF PROPOSITION 6

We consider the probability of promotion; if the system is in a period t and the probability of an A-agent ($\mathcal{P}_A(prom)$) being promoted is greater than that for a B-agent ($\mathcal{P}_B(prom)$), then the proportion of A -agents in period $t+1$ is greater than in t. Therefore, the population of A-agents has increased and the population of B-agents has decreased.

An agent can be promoted in two different events, i.e., he is either eligible and successful ($E.S.$) or he is eligible and but unsuccessful ($E.U.$). Therefore, the probability of an agent's following the strategy i being promoted (in period t) can be written as:

$$\mathcal{P}_{i,t}(prom) = \mathcal{P}_t(prom/E.S.)\mathcal{P}_{i,t}(E.S.) + \mathcal{P}_t(prom/E.U.)\mathcal{P}_{i,t}(E.U.)$$

Thus, the probability of promotion for an agent who follows strategy i is a linear combination of the agent's probability of being in event $E.S.$ (i.e. $\mathcal{P}_{i,t}(E.S.)$) or in $E.U.$ (i.e. $\mathcal{P}_{i,t}(E.U.)$), where the weights are given by the probability of being promoted in each either event (i.e., $\mathcal{P}_t(prom/E.S.)$ and $\mathcal{P}_t(prom/E.U.)$).

Observe the Figure 12, remember the Figure 2 and 3 of section 3.1:

It is straightforward to derive the following probabilities:

$$\mathcal{P}_t(prom/E.S.) = \min\{\frac{\frac{k}{n}\theta}{ES_t}, 1\} \qquad \mathcal{P}_t(prom/E.U.) = \max\{\frac{\theta\frac{k}{n} - ES_t}{EU_t}, 0\}$$
$$\mathcal{P}_{A,t}(E.S.) = \frac{1}{a_t}ES_t^a \qquad \mathcal{P}_{A,t}(E.U.) = \frac{1}{a_t}EU_t^a$$
$$\mathcal{P}_{B,t}(E.S.) = \frac{1}{(1-a_t)}(ES_t - ES_t^a) \qquad \mathcal{P}_{B,t}(E.U.) = \frac{1}{(1-a_t)}(EU_t - EU_t^a)$$

Note that, $ES_t^a = a_t\mathcal{P}_{A,t}(E.S.)$, thus $\mathcal{P}_{A,t}(E.S.) = \frac{1}{a_t}ES_t^a$, we work in the same way to derive $\mathcal{P}_{i,t}(E.U.)$ and $\mathcal{P}_{i,t}(E.S.)$.

We want study the local stability of the system in $a = 1$ and $a = 0$.

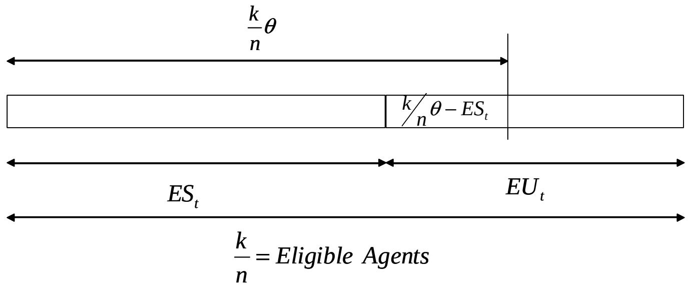


Figure 12:

For example, if we want to study the stability[45] of $a = 1$, then it is sufficient to study which promotion probability is greater when the state of the system is very close to one ($a_t \simeq 1$). If the promotion probability of an A-agent is greater(smaller) than that of a B-agent and that is true as much as the system closes to one is, then the state $a = 1$ will be locally stable(unstable).(Analogously to $a = 0$).

Notice that,

- if $a_t \simeq 1$ then:

$$\begin{aligned} ES_t &\simeq ES_t^a \simeq \frac{k}{n}P_1 \\ EU_t &\simeq EU_t^a \simeq \frac{k}{n}P_2 \end{aligned}$$

Therefore:

$$\begin{array}{cc} \mathcal{P}_t(prom/E.S.) \simeq \min\{\dfrac{\theta}{P_1}, 1\} & \mathcal{P}_t(prom/E.U.) \simeq \max\{\dfrac{\theta - P_1}{P_2}, 0\} \\ \mathcal{P}_{A,t}(E.S.) \simeq \frac{k}{n}P_1 & \mathcal{P}_{A,t}(E.U.) \simeq \dfrac{k}{n}P_2 \\ \mathcal{P}_{B,t}(E.S.) \simeq P_2 & \mathcal{P}_{B,t}(E.U.) \simeq 0 \end{array}$$

The probability $\mathcal{P}_{B,t}(E.S.)$ is straightforward to derive. We first we simplify[46] the expression $\dfrac{1}{(1 - a_t)}(ES_t - ES_t^a)$, afterwards, we calculate the limit of this expression when a_t goes to 1, and this limit is equal to P_2. We work analogously to derive $\mathcal{P}_{B,t}(E.U.)$.

The promotion probability for an A-agent when the state is close to one,$a_t \simeq 1$,will be:

$$\begin{gathered} \mathcal{P}_{A,a_t \simeq 1}(prom) = \min\{\frac{\theta}{P_1}, 1\}\tfrac{k}{n}P_1 + \max\{\frac{\theta - P_1}{P_2}, 0\}\tfrac{k}{n}P_2 \\ \mathcal{P}_{A,a_t \simeq 1}(prom) = \min\{\theta, P_1\}\tfrac{k}{n} + \max\{\theta - P_1, 0\}\tfrac{k}{n} \\ \mathcal{P}_{A,a_t \simeq 1}(prom) = \tfrac{k}{n}(\min\{\theta, P_1\} + \max\{\theta - P_1, 0\}) \\ \text{if } \theta < P_1 \text{ then } \mathcal{P}_{A,a_t \simeq 1}(prom) = \tfrac{k}{n}\theta \\ \text{if } \theta \geq P_1 \text{ then } \mathcal{P}_{A,a_t \simeq 1}(prom) = \tfrac{k}{n}(P_1 + \theta - P_1) = \tfrac{k}{n}\theta \end{gathered}$$

And for a B-agent:

[45] In this proof, the stability we refer to is always the local stability

[46] See Section 5 to for an explicit expression of ES_t and ES_t^a.

$$\mathcal{P}_{B,a_t\simeq 1}(prom) = \min\{\frac{\theta}{P_1},1\}P_2 + \max\{\frac{\theta-P_1}{P_2},0\}0$$
$$\mathcal{P}_{B,a_t\simeq 1}(prom) = \min\{\frac{\theta}{P_1},1\}P_2$$
$$\text{if } \theta < P_1 \text{ then } \mathcal{P}_{B,a_t\simeq 1}(prom) = \frac{\theta}{P_1}P_2$$
$$\text{if } \theta \geq P_1 \text{ then } \mathcal{P}_{B,a_t\simeq 1}(prom) = P_2$$

Thus;

$$\begin{aligned}
\mathcal{P}_{A,a_t\simeq 1}(prom) &\simeq \frac{k}{n}\theta \\
\mathcal{P}_{B,a_t\simeq 1}(prom) &\simeq \begin{cases} \frac{\theta}{P_1}P_2 \text{ if } \theta < P_1 \\ P_2 \text{ if } \theta \geq P_1 \end{cases}
\end{aligned}$$

We know that if $\mathcal{P}_{A,a_t\simeq 1}(prom) \geq (<)\mathcal{P}_{B,a_t\simeq 1}(prom)$ then $a = 1$ is stable (unstable).

Thus, if $\theta < P_1$ then $a = 1$ will be stable (unstable) provided that $\frac{k}{n}\theta \geq (<) \frac{\theta}{P_1}P_2 \Leftrightarrow \frac{k}{n} \geq (<) \frac{P_2}{P_1}$.

On the other hand, if $\theta \geq P_1$ then $a = 1$ will be stable (unstable) provided that $\frac{k}{n}\theta \geq (<)P_2$.

Analogously, we can study the state $a = 0$ and we can derive the promotion probabilities in this case:

$$\begin{aligned}
\mathcal{P}_{B,a_t\simeq 0}(prom) &\simeq \frac{k}{n}\theta \\
\mathcal{P}_{A,a_t\simeq 0}(prom) &\simeq \begin{cases} \frac{\theta}{P_2}P_1 \text{ if } \theta < P_2 \\ P_1 \text{ if } \theta \geq P_2 \end{cases}
\end{aligned}$$

Thus, if $\theta < P_2$ then $a = 0$ will be stable (unstable) provided that $\frac{k}{n}\theta \geq (<) \frac{\theta}{P_2}P_1 \Leftrightarrow \frac{k}{n} \geq (<) \frac{P_1}{P_2}$. Since $P_1 > P_2$, the state $a = 0$ will be unstable provided that $\theta < P_2$.

On the other hand, if $\theta \geq P_2$ then $a = 0$ will be stable (unstable) provided that $\frac{k}{n}\theta \geq (<)P_1$.

Therefore, if $\frac{k}{n}\theta < P_1$ then the state $a = 0$ will be unstable, and if $\frac{k}{n}\theta \geq P_1$ that state will be stable (since if $\frac{k}{n}\theta \geq P_1 \Rightarrow \theta \geq \frac{n}{k}P_1 > P_1 > P_2$).

In short:

if $\theta < P_1$ then $\begin{cases} a = 1 \text{ stable if } \frac{k}{n} \geq \frac{P_2}{P_1} \\ a = 1 \text{ unstable if } \frac{k}{n} < \frac{P_2}{P_1} \end{cases}$	if $\theta < P_2$ then $a = 0$ unstable
if $\theta \geq P_1$ then $\begin{cases} a = 1 \text{ stable if } \frac{k}{n}\theta \geq P_2 \\ a = 1 \text{ unstable if } \frac{k}{n}\theta < P_2 \end{cases}$	if $\theta \geq P_2$ then $\begin{cases} a = 0 \text{ stable if } \frac{k}{n}\theta \geq P_1 \\ a = 0 \text{ unstable if } \frac{k}{n}\theta < P_1 \end{cases}$

Note that, if $\theta < P_1$ then $\frac{k}{n}\theta < P_1$ and $a = 0$ will be unstable.

Since $P_1 > P_2$ we can state;

- If $\theta < P_1$ and,

  $\frac{k}{n} < \frac{P_2}{P_1}$ then $a = 0$ and $a = 1$ are unstable.

  $\frac{k}{n} \geq \frac{P_2}{P_1}$ then $a = 0$ is unstable and $a = 1$ is stable.

- If $\theta \geq P_1$ and,
  $\frac{k}{n}\theta < P_2$ then $a = 0$ and $a = 1$ are unstable.
  $P_2 \leq \frac{k}{n}\theta \leq P_1$ then $a = 0$ is unstable and $a = 1$ is stable.
  $\frac{k}{n}\theta > P_1$ then $a = 0$ and $a = 1$ are stable. ■

PROOF OF PROPOSITION 7
The solution of our maximization problem is

$$\bar{a} = \arg\max_{a \in [0,1]} \Pi(a)$$

The payoff function is:

$$\Pi(a) = \frac{1}{2}(P_2(\eta + \sigma) - \eta) + a(P_1(1 + \eta) + P_2(1 - \sigma)) - a^2\frac{1}{2}(P_1 + P_2)(2 + \eta - \sigma)$$

The first and second derivatives are:

$$\Pi_a = (P_1(1 + \eta) + P_2(1 - \sigma)) - a(P_1 + P_2)(2 + \eta - \sigma)$$

$$\Pi_{aa} = (P_1 + P_2)(2 + \eta - \sigma)$$

The payoff function is concave if $\sigma < 2 + \eta$ convex if $\sigma > 2 + \eta$ and lineal if $\sigma = 2 + \eta$.

Suppose that the payoff function is either convex or lineal ($\sigma \geq 2 + \eta$). We would then have just two candidates for the solution of our maximization problem, which would be $a = 1$ and $a = 0$.

We have that $\Pi(a = 0) = \frac{1}{2}(P_2(\eta + \sigma) - \eta)$ and $\Pi(a = 1) = \frac{1}{2}(P_1(\eta + \sigma) - \eta)$ if $P_1 > P_2$ then $\bar{a} = 1$ and if $P_2 > P_1$ then $\bar{a} = 0$.

If we suppose that the payoff function is concave ($\sigma < 2 + \eta$), then the maximum will be:

$$\hat{a} = \frac{P_1(1 + \eta) + P_2(1 - \sigma)}{(P_1 + P_2)(2 + \eta - \sigma)}$$

Since $a \in [0, 1]$, if $\hat{a} > 1$ (< 0), then the solution to our maximization problem is $\bar{a} = 1$ ($= 0$). As $\sigma < 2 + \eta$, the denominator is positive and it is straightforward to show that:

$$\hat{a} > 0 \Leftrightarrow \sigma < 1 + \frac{P_1}{P_2}(1 + \eta)$$

$$\hat{a} > 1 \Leftrightarrow \sigma < 1 + \frac{P_2}{P_1}(1 + \eta)$$

If $P_1 > P_2$ then $1 + \frac{P_2}{P_1}(1 + \eta) < 1 + \frac{P_1}{P_2}(1 + \eta)$ (Note that in this case $1 + \frac{P_2}{P_1}(1 + \eta) < 2 + \eta$)

If $P_2 > P_1$ then $1 + \frac{P_1}{P_2}(1 + \eta) < 1 + \frac{P_2}{P_1}(1 + \eta)$ (Note that in this case $1 + \frac{P_1}{P_2}(1 + \eta) < 2 + \eta$)

Therefore if $P_1 > P_2$ then $\bar{a} = \begin{cases} \frac{P_1(1+\eta)+P_2(1-\sigma)}{(P_1+P_2)(2+\eta-\sigma)} & \text{if} \quad \sigma < 1 + \frac{P_2}{P_1}(1 + \eta) \\ 1 & \text{if} \quad \sigma \geq 1 + \frac{P_2}{P_1}(1 + \eta) \end{cases}$ and if $P_2 > P_1$ then $\bar{a} = \begin{cases} \frac{P_1(1+\eta)+P_2(1-\sigma)}{(P_1+P_2)(2+\eta-\sigma)} & \text{if} \quad \sigma < 1 + \frac{P_1}{P_2}(1 + \eta) \\ 0 & \text{if} \quad \sigma \geq 1 + \frac{P_1}{P_2}(1 + \eta) \end{cases}$ ■

REFERENCES

Elaydi, Saber . "An introduction to difference equations" Springer-Verlag New York,(1996)

Harrington,J. E. (1998a). "The Social Selection of Flexible and Rigid Agents." American Economic Review 88: 63-82.

Harrington,J. E. (1998b). "Progressive Ambition, Electoral Selection, and the Creation of Ideologies" Economics of Governance, Vol. 1, Issue 1, 2000

Harrington,J. E. (1999a). "Rigidity of Social Systems" Journal of Political Economy, 107 40-64.

Harrington,J. E. (1999b). "Rigidity and Flexibility in Social Systems: Transitional Dynamics" in Computational Models in Political Economy, K. Kollman, J. Miller, and S. Page, eds., The MIT Press, 2003

Hvide,H. K.and Kristiansen, E. G (2002). "Risk taking in selection contests" Games and Economic Behavior 42, 172-179

Lazear, E., Rosen, S., (1981). "Rank-order tournaments as optimum labor contracts". Journal of Political Economy 89, 841–864.

Rosen, S., 1986. Prizes and incentives in elimination tournaments. American Economic Review. 76, 701–715.

Vega-Redondo, Fernando. "Unfolding Social Hierarchies", Journal of Economic Theory 90, 177-203 (2000)